

Luís Eduardo Abatti

**REGULAÇÃO DA EXPRESSÃO DE *SOX2* EM CÉLULAS-
TRONCO EMBRIONÁRIAS E CÉLULAS-TRONCO INDUZIDAS
A EPIBLASTO DE CAMUNDONGO**

Trabalho de Conclusão de Curso
apresentado à Coordenadoria do Curso
de Graduação em Ciências Biológicas
da Universidade Federal de Santa
Catarina como requisito parcial para a
conclusão do currículo do curso em
Ciências Biológicas.

Orientador: Prof. Dr. Ricardo
Castilho Garcez (Universidade
Federal de Santa Catarina).
Coorientadora: Prof.^a Dr.^a Jennifer
A. Mitchell (Universidade de
Toronto).

Florianópolis
2015

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Abatti, Luís Eduardo
REGULAÇÃO DA EXPRESSÃO DE SOX2 EM CÉLULAS-TRONCO
EMBRIONÁRIAS E CÉLULAS-TRONCO INDUZIDAS A EPIBLASTO DE
CAMUNDONGO / Luís Eduardo Abatti ; orientador, Ricardo
Castilho Garcez ; coorientadora, Jennifer A. Mitchell. -
Florianópolis, SC, 2015.
96 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro de Ciências
Biológicas. Graduação em Ciências Biológicas.

Inclui referências

1. Ciências Biológicas. 2. Controle de expressão gênica.
3. Regiões reguladoras. 4. Diferenciação celular. 5.
Células-tronco. I. Garcez, Ricardo Castilho. II. Mitchell,
Jennifer A.. III. Universidade Federal de Santa Catarina.
Graduação em Ciências Biológicas. IV. Título.

Luís Eduardo Abatti

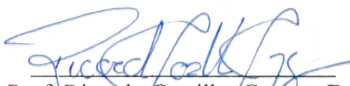
**REGULAÇÃO DA EXPRESSÃO DE *SOX2* EM CÉLULAS-
TRONCO EMBRIONÁRIAS E CÉLULAS-TRONCO INDUZIDAS
A EPIBLASTO DE CAMUNDONGO**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de Bacharel em Ciências Biológicas, e aprovado em sua forma final pelo Programa de Ciências Biológicas da Universidade Federal de Santa Catarina.

Florianópolis, 01 de dezembro de 2015.

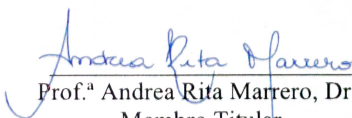
Prof.^a Maria Risoleta Freire Marques, Dr.^a
Coordenadora do Curso

Banca Examinadora:



Prof. Ricardo Castilho Garez, Dr.
Orientador

Universidade Federal de Santa Catarina



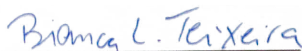
Prof.^a Andrea Rita Marrero, Dr.^a
Membro Titular

Universidade Federal de Santa Catarina



Prof. Geison de Souza Izídio, Dr.
Membro Titular

Universidade Federal de Santa Catarina



Prof.^a Bianca Luise Teixeira, Dr.^a
Membro Suplente

Universidade Federal de Santa Catarina

Dedico esse trabalho aos meus pais,
Regina Cavalli e Francisco Antônio
Abatti, pelo amor incondicional. E aos
meus amigos, pelo carinho e amizade.

AGRADECIMENTOS

Agradeço a todos que participaram e contribuíram de alguma forma em minha vida acadêmica, em especial nos últimos anos que precederam a elaboração desse trabalho de conclusão de curso.

À minha mãe, Regina Cavalli, e ao meu pai, Francisco Antônio Abatti, por todo seu amor, dedicação e apoio durante a minha criação e os meus anos de graduação. Obrigado por acreditarem em meus sonhos e me apoiarem incondicionalmente na busca dos meus objetivos;

Aos meus melhores amigos, Artur Felipe Figueira e Paulo Roberto Lopes Junior, pelo grande carinho, participação e apoio moral, físico e emocional. Os últimos anos teriam sido muito mais difíceis sem vocês;

Ao Governo Federal, Ministério da Educação e CNPq, por possibilitarem a realização do meu intercâmbio para o Canadá, através do programa Ciência sem Fronteiras. Nenhuma parte desse trabalho teria sido realizada sem a concessão dessa grande oportunidade;

À Professora Jennifer Mitchell, assim como toda a equipe de seu laboratório, por me acolherem e me guiarem durante meu estágio de pesquisa na Universidade de Toronto. Obrigado por receberem um estudante estrangeiro e abrirem portas para a realização desse projeto;

Às minhas amigas canadenses Hannah Samuels e Helena Zhang pelos grandes aprendizados, momentos inesquecíveis e indispensável companhia. Minha estadia no Canadá não teria sido tão proveitosa sem vocês;

Ao Professor Ricardo Castilho Garcez, pela grande ajuda e disponibilidade na elaboração, escrita, revisão e apresentação desse trabalho;

Aos meus colegas da Universidade Federal de Santa Catarina, especialmente Carolina Erbes, Fernanda Lamin Henrique e Gabriela Alanis. É difícil escolher os melhores momentos que vivenciamos entre realização de trabalhos, estudos ou mesmo companhia durante as aulas. Esses cinco anos teriam sido monótonos sem vocês;

Aos diversos professores que marcaram minha graduação e que fizeram-me apaixonar pelo curso de Ciências Biológicas, especialmente pela área de pesquisa molecular. Espero um dia chegar ao nível de conhecimento que cada um de vocês possui e ter o grande prazer de compartilhá-lo com outras pessoas, do mesmo modo que vocês o fazem.

RESUMO

Células-tronco embrionárias são células pluripotentes extraídas da massa interna do blastocisto. Elas darão origem a todos os tecidos do corpo e anexos embrionários. Após a implantação na parede uterina, as células da massa interna do blastocisto diferenciam-se em epiblasto e hipoblasto. As células do epiblasto ainda são capazes de originar derivados dos três folhetos embrionários, mas perdem a capacidade de originar anexos embrionários. O fator de transcrição Sox2 é um dos principais reguladores da pluripotencialidade celular. Quatro regiões reguladoras foram descritas no controle da expressão do gene *Sox2*, SRR1, SRR2, SRR18 e SCR. Porém, não é evidente se todas essas regiões reguladoras demonstram similar importância em seu controle. Para elucidar essa questão, o presente trabalho analisou os efeitos da deleção das regiões reguladoras SRR1, SRR2, SRR18 e SCR sobre a expressão de *Sox2* em células-tronco embrionárias em modelo murino. Somente a deleção da região SCR, em ambos os alelos, eliminou a expressão da proteína Sox2. A seguir, células-tronco embrionárias contendo deleções heterozigóticas combinadas nas regiões SRR1, SRR2, SRR18, simultaneamente, e na região SCR foram diferenciadas em células-tronco induzidas a epiblasto. A expressão de genes indicadores de estado celular foi analisada antes e após o processo de diferenciação. Enquanto a transcrição dos genes *Esrrb*, *Klf4*, *Nanog*, *Oct4* e *Tbx3* diminuiu, a transcrição dos genes *Dnmt3a*, *Otx2*, *Pou3f1* e *Wnt8a* aumentou após a diferenciação para células induzidas a epiblasto. A expressão de *Sox2* proveniente de cada alelo também foi analisada. Novamente, somente a deleção da região SCR resultou em grande impacto na expressão de *Sox2* em ambos os estados celulares. Por fim, a região SCR foi caracterizada *in silico* como tendo uma grande quantidade de modificações em caudas de histonas que permitem a aderência de fatores de transcrição em células-tronco embrionárias. Muitos deles contribuem para a sustentação da pluripotencialidade de células-tronco embrionárias, como Ctf, E2f1, Esrrb, Klf4, Nanog, Oct4, Smad1, Stat3, Tcf21 e o próprio Sox2. A região SCR também possui diferentes motivos de DNA que permitem a ligação desses fatores de transcrição. Juntos, esses dados sugerem que a expressão de *Sox2* é controlada majoritariamente pela região reguladora SCR em células-tronco embrionárias e células-tronco induzidas a epiblasto.

Palavras-chave: Controle de expressão gênica. Regiões reguladoras. Diferenciação celular. Fatores de transcrição. Epigenética.

ABSTRACT

Embryonic stem cells are pluripotent cells derived from the inner cell mass of the blastocyst. These cells will originate all tissues from the body, including extra-embryonic tissues. After the implantation in the uterus wall, cells from the inner cell mass will differentiate into the epiblast and the hypoblast. At this time, the epiblast is still capable of establishing all three germ layers, but it loses its ability to form extra-embryonic tissues. Sox2 transcription factor is one of the main regulators of pluripotency in the cell. Four major enhancers have been established for controlling the *Sox2* gene, SRR1, SRR2, SRR18 and the SCR. However, it is unclear if these enhancers have the same influence on *Sox2* expression. To clarify this matter, the current study has analysed the effects of deleting enhancers SRR1, SRR2, SRR18 and SCR on *Sox2* expression in mouse embryonic stem cells. The SCR region has shown to be the only enhancer capable of eliminating Sox2 protein expression when deleted on both alleles. Next, embryonic stem cells containing combined heterozygous deletions on enhancers SRR1, SRR2, SRR18 together and on the SCR region alone were differentiated toward epiblast-like stem cells. Gene expression of cell state-specific genes was analyzed both before and after differentiation occurred. While *Esrrb*, *Klf4*, *Nanog*, *Oct4* and *Tbx3* were down-regulated, *Dnmt3a*, *Otx2*, *Pou3f1* and *Wnt8a* were up-regulated after epiblast-like differentiation. *Sox2* expression from each allele was also assessed. Once again, only the SCR deletion deeply impacted *Sox2* expression in both cellular states. Moreover, the SCR region has been characterized *in silico* as having a high amount of histone tail modifications that allow transcription factor binding in embryonic stem cells. Many of these transcription factors are related to the maintenance of embryonic stem cells pluripotency, such as Ctf, E2f1, Esrrb, Klf4, Nanog, Oct4, Smad1, Stat3, Tcf21 and Sox2. Furthermore, the SCR region contains different DNA motifs throughout its sequence that allow for the binding of these transcription factors. Together, these data suggest that the expression of *Sox2* is mainly controlled by the SCR region in embryonic stem cells and epiblast-like stem cells.

Keywords: Gene expression control. Enhancers. Cell differentiation. Transcription factors. Epigenetics.

LISTA DE FIGURAS

Figura 1: Desenvolvimento embrionário inicial de camundongos.....	32
Figura 2: Mapa genômico da posição das regiões reguladoras SRR1, SRR2, SRR18 e SCR (SRR104 a SRR112) em relação ao gene alvo <i>Sox2</i> no cromossomo 3 de camundongos	37
Figura 3: Síntese dos procedimentos utilizado para a preparação das células-tronco embrionárias utilizadas durante o primeiro bloco experimental.....	40
Figura 4: Síntese dos procedimentos utilizados para a preparação das células-tronco embrionárias e células-tronco induzidas a epiblasto utilizadas durante o segundo bloco experimental	43
Figura 5: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 30, 54 e 91), $\Delta\text{SRR1}/2/18^{\text{Cast}}$ (Clone 39) e $\Delta\text{SRR2-SRR104}^{129/\text{Cast}}$ (Clone 18 e 21)	51
Figura 6: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 1, Clone 4, Clone 30 e Clone 37), $\Delta\text{SRR18-104}^{\text{Cast}}$ (Clone 21 e 39) e $\Delta\text{SRR18-104}^{129/\text{Cast}}$ (Clone 70).....	51
Figura 7: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), F1 (Diferenciado), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18-104}^{\text{Cast}}$ (Clone 10), $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 62), $\Delta\text{SOX2}^{\text{Cast}}$ (Clone 36), $\Delta\text{SRR1}^{129/\text{Cast}} + \Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 25 e 39) e $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 36)	53
Figura 8: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 30) e $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 62, 54 e 17)	54
Figura 9: Imagem de microscópio de luz invertida dos grupos celulares F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1}/2/18^{\text{Cast}}$ e de seus respectivos tratamentos, controle e induzido a epiblasto, após 48 horas de cultivo em seus meios de cultura N2B27 e N2B27 indutor de epiblasto	56
Figura 10: Expressão do gene <i>Gapdh</i> obtida através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1}/2/18^{\text{Cast}}$	57
Figura 11: Expressão dos genes <i>Esrrb</i> , <i>Klf4</i> , <i>Nanog</i> , <i>Oct4</i> e <i>Tbx3</i> obtidas através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1}/2/18^{\text{Cast}}$	58
Figura 12: Expressão dos genes <i>Dnmt3a</i> , <i>Otx2</i> , <i>Pou3fl</i> e <i>Wnt8a</i> obtidas através de PCR quantitativo de três réplicas biológicas (n=3) nos	

tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$	60
Figura 13: Expressão do gene <i>Sox2</i> proveniente do alelo <i>I29</i> e do alelo <i>Castaneus</i> obtida através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$	62
Figura 14: Mapa genômico contendo informações das regiões reguladoras SRR1, SRR2, SRR18 e SCR	64
Figura 15: Mapa genômico contendo dez fatores de transcrição com atividade de ligação nas áreas reguladoras SRR1, SRR2, SRR18 e SCR	66
Figura 16: Possíveis locais de ligação dos fatores de transcrição <i>Esrrb</i> , <i>Klf4</i> , <i>Sox2</i> , <i>Stat3</i> e <i>Tfcp2l1</i> na região SCR conforme resultados obtidos <i>in silico</i> pelo software Jaspasr	67
Figura 17: Mecanismos de controle da expressão de <i>Sox2</i> em células-tronco embrionárias	77

LISTA DE QUADROS

Quadro 1: Relação da nomenclatura de deleções realizadas em células-tronco embrionárias através do sistema CRISPR/Cas9, clones celulares gerados, alelos deletados e regiões reguladoras afetadas por cada deleção.	38
Quadro 2: Sequência de <i>primers</i> utilizadas para os experimentos de PCR quantitativo de transcrição reversa (RT-qPCR) desenhados através do uso do Primer3Plus	45

LISTA DE ABREVIATURAS E SIGLAS

129 -	<i>129 Allele</i>	Alelo 129
Cas9 -	<i>CRISPR associated protein 9</i>	Proteína associada ao CRISPR 9
Cast -	<i>Castaneus Allele</i>	Alelo <i>Castaneus</i>
ChIP-Seq -	<i>Chromatin immunoprecipitation Sequencing</i>	Sequenciamento de imunoprecipitação de cromatina
CHIR99021 -	<i>Glycogen synthase kinase 3 inhibitor</i>	Inibidor da enzima glicogênio sintase cinase 3
CRISPR -	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i>	Repetições palindrômicas curtas agrupadas e regularmente interespaçadas
Ctcf -	<i>CCCTC-binding factor</i>	Fator de ligação <i>CCCTC</i>
DMEM -	<i>Dulbecco's Modified Eagle's Medium</i>	Meio de cultura de Eagle modificado por Dulbecco
DNA -	<i>Deoxyribonucleic acid</i>	Ácido desoxirribonucleico
Dnmt3a* -	<i>DNA (cytosine-5)-methyltransferase 3A</i>	DNA (citosina-5)-metiltransferase 3A
E2f1* -	<i>E2F transcription factor 1</i>	Fator de transcrição E2F1
EDTA -	<i>Ethylenediamine tetraacetic acid</i>	Ácido etilenodiamino tetra-acético
Esrrb* -	<i>Estrogen-Related Receptor Beta</i>	Receptor relacionado ao estrogênio beta
F12 -	<i>Nutrient Mixture F12</i>	Mistura de nutrientes F12
Fgf2 -	<i>Fibroblast growth factor 2</i>	Fator de crescimento de fibroblasto 2
Gapdh* -	<i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>	Gliceraldeído-3-Fosfato Desidrogenase
H3K27ac -	<i>H3 Lysine 27 acetylation</i>	Acetilação na lisina 27 da cauda de histonas H3
H3K4me1 -	<i>H3 Lysine 4 methylation</i>	Metilação na lisina 4 da cauda de histonas H3
H3K4me3 -	<i>H3 Lysine 4 trimethylation</i>	Trimetilação na lisina 4 da cauda de histonas H3
H3K9ac -	<i>H3 Lysine 9 acetylation</i>	Acetilação na lisina 9 da cauda de histonas H3
Klf4* -	<i>Kruppel-like factor 4</i>	Fator do tipo Kruppel 4
LIF -	<i>Leukemia inhibitory factor</i>	Fator inibidor de leucemia
MEM -	<i>Minimum Essential Media</i>	Meio de cultura essencial

Mm9 /	<i>Mus musculus genome</i>	Genoma de <i>Mus musculus</i>
Mm10 -	<i>version 9 / 10</i>	versão 9 / 10
<i>Nanog</i>* -	<i>Nanog Homeobox</i>	Segmento homeótico Nanog
<i>Oct4</i>* -	<i>Octamer-binding transcription factor 4</i>	Fator de transcrição ligador de octâmero 4
<i>Otx2</i>* -	<i>Orthodenticle Homeobox 2</i>	Segmento homeótico ortodentilo 2
PBS -	<i>Phosphate buffered saline</i>	Tampão fosfato-salino
PCR -	<i>Polymerase Chain Reaction</i>	Reação em cadeia da polimerase
PD0325901 -	<i>Mitogen-activated protein kinase kinase inhibitor</i>	Inibidor da enzima cinase ativada por mitógenos
<i>Pou3f1</i>* -	<i>POU Class 3 Homeobox 1</i>	Segmento homeótico POU classe 3
PVDF -	<i>Polyvinylidene fluoride</i>	Fluoreto de polivinilideno
RIPA -	<i>Radioimmunoprecipitation Assay</i>	Ensaio de rádio imunoprecipitação
RNA -	<i>Ribonucleic acid</i>	Ácido ribonucleico
RT-qPCR -	<i>Reverse transcription polymerase chain reaction</i>	PCR quantitativo de transcrição reversa
SCR -	<i>Sox2 Control Region</i>	Região controle do gene <i>Sox2</i>
<i>Smad1</i>* -	<i>SMAD family member 1</i>	Membro da família SMAD 1
<i>Sox2</i>* -	<i>Sex determining region Y-box 2</i>	Região determinadora de sexo Y-2
SRR -	<i>Sox2 Regulatory Region</i>	Região reguladora de <i>Sox2</i>
<i>Stat3</i>* -	<i>Signal Transducer and Activator of Transcription 3</i>	Transdutor de sinal e ativador de transcrição 3
TBS -	<i>Tris-buffered saline</i>	Tampão salina tris
<i>Tbx3</i>* -	<i>T-Box 3</i>	Região T-3
<i>Tfcp2l1</i>* -	<i>Transcription Factor CP2-Like 1</i>	Fator de transcrição do tipo CP2 1
TTBS -	<i>Tween Tris-buffered saline</i>	Tampão salina tris tween
UCSC -	<i>University of California, Santa Cruz</i>	Universidade da Califórnia, Santa Cruz
<i>Wnt8a</i>* -	<i>Wingless-Type MMTV Integration Site Family, Member 8A</i>	Membro 8A do sítio integrador MMTV do tipo âptero

* A nomenclatura de genes e proteínas adotada nesse trabalho segue as orientações do Comitê Internacional de Nomenclatura de Camundongos, onde genes são representados pela formatação em itálico (ex: *Sox2*) enquanto proteínas são apresentadas com formatação normal (ex: Sox2).

SUMÁRIO

1	INTRODUÇÃO.....	21
1.1	O DESENVOLVIMENTO DE CÉLULAS-TRONCO EMBRIONÁRIAS	21
1.2	A REGULAÇÃO DA EXPRESSÃO GÊNICA	23
1.2.1	Promotor	24
1.2.2	Fatores de transcrição	24
1.2.3	Regiões reguladoras	25
1.2.4	Modificações de histonas	26
1.3	O CONTROLE DA EXPRESSÃO DE <i>SOX2</i>	27
1.3.1	O sistema de deleção CRISPR/Cas9	28
1.3.2	As regiões reguladoras de <i>Sox2</i>.....	29
1.4	O PERFIL DE EXPRESSÃO GÊNICA DURANTE A TRANSIÇÃO PARA O EPIBLASTO	30
1.5	OS FATORES DE TRANSCRIÇÃO DENTRO DE REGIÕES REGULADORAS.....	32
1.6	JUSTIFICATIVA.....	33
1.7	HIPÓTESES.....	33
2	OBJETIVOS.....	35
2.1	OBJETIVO GERAL	35
2.2	OBJETIVOS ESPECÍFICOS.....	35
3	MATERIAL E MÉTODOS.....	37
3.1	PRIMEIRO BLOCO EXPERIMENTAL	37
3.1.1	O cultivo de células-tronco embrionárias	37
3.1.2	Extração e cálculo de concentração proteica.....	39
3.1.3	Detecção da proteína Sox2 em relação ao controle Gapdh ...	40
3.2	SEGUNDO BLOCO EXPERIMENTAL	42
3.2.1	O cultivo de células-tronco embrionárias	42
3.2.2	Indução do estado de epiblasto	43
3.2.3	Extração de RNA	43
3.2.4	Desenho de primers.....	44
3.2.5	Quantificação da expressão gênica	45
3.2.6	Análise estatística do perfil de expressão gênica	46
3.3	TERCEIRO BLOCO EXPERIMENTAL	46
3.3.1	O papel molecular da região SCR	46
3.3.2	Os fatores de transcrição relacionados à região SCR.....	47
4	RESULTADOS.....	49
4.1	PRIMEIRO BLOCO EXPERIMENTAL	49
4.1.1	Detecção da proteína Sox2 em diferentes clones celulares....	49
4.2	SEGUNDO BLOCO EXPERIMENTAL	55

4.2.1	A diferenciação de células-tronco embrionárias	55
4.2.2	O padrão de expressão gênica de células-tronco induzidas a epiblasto	56
4.3	TERCEIRO BLOCO EXPERIMENTAL	62
4.3.1	O papel molecular da região SCR	62
4.3.2	Os fatores de transcrição relacionados à região SCR.....	65
5	DISCUSSÃO	69
6	CONCLUSÕES	79
7	PERSPECTIVAS	81
	REFERÊNCIAS BIBLIOGRÁFICAS.....	83
	ANEXO A	96

1 INTRODUÇÃO

1.1 O DESENVOLVIMENTO DE CÉLULAS-TRONCO EMBRIONÁRIAS

Células-tronco embrionárias são células pluripotentes extraídas da massa interna de células do blastocisto, ou seja, do embrioblasto, formado entre o quarto ou quinto dia de desenvolvimento em humanos, e entre o terceiro e quarto dia em camundongos (EVANS; KAUFMAN, 1981; MARTIN, 1981; THOMSON *et al.*, 1998). O desenvolvimento embrionário começa após a fecundação. Nesse momento, o zigoto, resultante da união do espermatozoide e do ovócito secundário, inicia uma série de divisões, formando os blastômeros. A cada clivagem, esses blastômeros se tornam menores, até que formem uma esfera compacta de células durante a fase conhecida como compactação. Essa esfera maciça é chamada de mórula, e contém cerca de 12 a 32 células embrionárias. Quando a mórula alcança a área uterina, há a formação de uma cavidade interna cheia de fluido, a blastocele, e a massa de células passa a se chamar, então, blastocisto (MOORE; PERSAUD, 2008).

O blastocisto é constituído por duas estruturas celulares distintas. A parte externa, chamada de trofoblasto, dará origem ao componente embrionário da placenta. Já a parte interna, chamada de embrioblasto ou massa celular interna, dará origem ao embrião (MOORE; PERSAUD, 2008). É nesse momento que células-tronco embrionárias podem ser extraídas da massa celular interna para cultivo *in vitro*. Por serem responsáveis pela formação completa do embrião, essas células possuem o potencial de gerar derivados de todas as três camadas germinativas: ectoderme, mesoderme e endoderme (NIWA, 2007; POH *et al.*, 2014). Assim sendo, elas possuem elevada capacidade de autorrenovação – a possibilidade de proliferar, gerando células idênticas à original – e de diferenciação, podendo gerar células especializadas de diferentes tecidos.

Até o presente estágio de blastocisto, o destino embrionário de cada célula pode ser alterado caso elas tenham sua posição modificada dentro do embrião em formação (HILLMAN; SHERMAN; GRAHAM, 1972). Além disso, cada célula extraída do embrioblasto, quando reimplantada em um outro blastocisto em formação, possui a capacidade de contribuir para a formação completa desse novo embrião, formando uma químera (GARDNER, 1998). Isto é, cada célula embrionária, até esse ponto, tem a capacidade igualitária de gerar todo o embrião.

Diversos genes são conhecidos por caracterizarem o estado não-diferenciado de células-tronco embrionárias. Dentre eles, os genes *Klf4*, *Nanog*, *Oct4* e *Sox2* são altamente expressos nesse estado, sendo que a inibição ou remoção de qualquer um desses genes resulta na inevitável

diferenciação dessas células-tronco embrionárias. Por conta disso, os fatores de transcrição, produtos da expressão desses genes, parecem ter um papel fundamental na manutenção da pluripotencialidade de células-tronco (CHAMBERS; TOMLINSON, 2009). Além disso, Takahashi e Yamanaka (2006) demonstraram que a reintrodução dos fatores de transcrição Oct4, Sox2 e Klf4 em células maduras e diferenciadas reverteram seu estado celular para uma condição similar a de células embrionárias, dando origem às células-tronco de pluripotência induzida. Desse modo, esses fatores são conhecidos como marcadores de pluripotencialidade (TAKAHASHI; YAMANAKA, 2006).

Conforme o embrião se desenvolve, o ilimitado potencial de proliferação celular da massa interna é reduzido devido à diferenciação e divisão do embrioblasto. Esse evento ocorre no período em que o blastocisto se implanta no útero da mãe, cerca de 8 dias após a fecundação em humanos, e em cerca de 5 dias em camundongos (WILCOX; BAIRD; WEINBERG, 1999; DEB; REESE; PARIA, 2006). A camada de células voltada para a superfície da blastocele formará o hipoblasto, ao mesmo tempo que a camada celular voltada para a parede uterina gerará o epiblasto (NIWA, 2007). Enquanto o epiblasto dará origem a todos os tecidos do embrião, as células do hipoblasto participarão do desenvolvimento de tecidos extra-embrionários, como a placenta (WOLPERT, 2008). Como consequência dessa divisão, as células do epiblasto terão seu perfil celular modificado, tanto morfológicamente quanto molecularmente, pois perderão a capacidade de gerar um organismo inteiro com todas suas estruturas extra-embrionárias necessárias para o desenvolvimento (ROSSANT, 2008). Além disso, células do epiblasto perdem a capacidade de contribuir para a formação de um novo organismo na forma de uma quimera, quando injetadas em um blastocisto em desenvolvimento (TESAR *et al.*, 2007).

A transição entre o embrioblasto e o epiblasto pode ser reproduzida, *in vitro*, através da exposição de células-tronco embrionárias a Activin A e Fgf2 – importantes fatores de crescimento para a manutenção de células do epiblasto (Greber *et al.*, 2010). Esses fatores irão induzir as células do embrioblasto a entrarem em um estado celular semelhante ao epiblasto (*epiblast-like cell state*, ou, células-tronco induzidas a epiblasto). Nesse novo estado, muitos dos fatores pluricelulares fundamentais, incluindo *Klf4*, *Nanog*, *Oct4* e *Sox2*, ainda serão expressos, porém em diferentes níveis. Dessa forma, essas células induzidas a epiblasto fornecem um excelente modelo de estudo dos mecanismos moleculares que controlam a transição entre o embrioblasto e o epiblasto, ou seja, entre um estado totalmente indiferenciado para um

estado com menor capacidade de pluripotência. Esses mecanismos moleculares irão, de forma similar, governar a transição para outros tipos celulares mais tardiamente (HAYASHI *et al.*, 2011; BUECKER *et al.*, 2014).

1.2 A REGULAÇÃO DA EXPRESSÃO GÊNICA

A regulação da expressão gênica é um processo fundamental não somente para o embrião em desenvolvimento, mas para todas as células. É através desse mecanismo que a célula consegue responder apropriadamente aos estímulos internos e externos, gerando uma resposta apropriada e diferenciada para cada situação, a partir de um mesmo material genético. Dentre os diferentes mecanismos que possibilitam essa modulação da resposta celular, o controle da transcrição parece ser o mais importante para a maioria das células, já que ele controla diretamente a síntese de moléculas de RNA, intermediárias entre o DNA e as proteínas (ALBERTS *et al.*, 2010).

O controle da transcrição de RNA mensageiro na célula eucariótica envolve, sumariamente, três elementos principais (GRIFFITHS *et al.*, 2009):

- Promotor: região responsável por permitir a ancoragem da RNA Polimerase II e o início da transcrição do DNA;
- Fatores de transcrição: proteínas que se ligam à RNA polimerase II e podem tanto estimular níveis mais eficientes de transcrição, quanto reprimi-la;
- Regiões reguladoras, ou *enhancers*: regiões do DNA que atraem e ligam fatores de transcrição, induzindo seu contato com o complexo da RNA polimerase II, seja por aproximação física e/ou por modificações na cromatina.

Esses três elementos são, ainda, sujeitos ao nível de compactação e acessibilidade do DNA, controlados por modificações epigenéticas nos elementos da cromatina, como as histonas (FELSENFELD *et al.*, 1996).

Para um melhor entendimento do funcionamento e ativação de certos genes, se faz necessário compreender como as regiões reguladoras os influenciam, bem como quais são os fatores de transcrição que estão envolvidos nesse processo. Isso permite que eventos como a diferenciação celular de células-tronco, a qual depende da ativação e desativação de genes em específico, possam ser melhor estudados e compreendidos.

1.2.1 Promotor

Promotores são regiões responsáveis por permitir a montagem e ligação do complexo de transcrição da RNA Polimerase II e o início da transcrição do gene alvo. Dentro dos eucariotos, um dos promotores mais conservados é a TATA box, uma sequência de nucleotídeos TATAAA à cerca de 25 ou 30 nucleotídeos acima do códon iniciador da transcrição (SMALE; KADONAGA, 2003). Dessa maneira, a TATA box age como um centro de ancoragem para a montagem e direcionamento do complexo de transcrição. Além disso, essa região também permite a interação de fatores de transcrição com o complexo de transcrição, criando mais um elemento de modulação da expressão gênica.

1.2.2 Fatores de transcrição

Fatores de transcrição são proteínas capazes de se ligar à certas sequências de DNA e estimular, ou bloquear, o recrutamento da enzima RNA Polimerase II para um gene alvo, e, portanto, permitir ou inibir a transcrição gênica.

Em geral, essas proteínas possuem dois domínios distintos. Um deles é responsável por determinar a ligação do fator de transcrição a uma certa região do DNA, enquanto o outro irá modular a taxa de transcrição. Esse último domínio irá interagir com o complexo de transcrição, constituído pela RNA polimerase II. Isso pode resultar no aumento da sua velocidade de formação, ou no aumento da sua eficiência de transcrição do gene alvo. Por outro lado, fatores de transcrição também podem bloquear a transcrição de certos genes. Esse fenômeno pode ocorrer através do bloqueio físico de ligação de outros ativadores na região gênica alvo, ou na formação de um complexo inibidor de ativação da RNA Polimerase II (LATCHMAN, 1997).

A atividade dos fatores de transcrição pode ser controlada de duas maneiras diferentes. Primeiramente, o controle da síntese de fatores de transcrição resulta em uma maior ou menor presença desses fatores no núcleo celular, o que, por consequência, pode aumentar ou diminuir sua capacidade de ação no DNA. Posteriormente, fatores de transcrição previamente sintetizados podem ser ativados conforme a demanda necessária no núcleo. Isso ocorre através de fosforilações, dissociações de proteínas inibitórias ou através de modificações químicas na sua estrutura proteica (MITCHELL; TJIAN, 1989). Assim, a célula consegue controlar a transcrição de seus genes através da expressão e ativação de fatores de transcrição, de modo a responder corretamente as demandas do ambiente e estado de desenvolvimento em que ela está inserida em determinado

momento. Portanto, cada estado celular possui um perfil de expressão de fatores de transcrição distinto (LATCHMAN, 1997).

Nem todos os fatores de transcrição se ligam somente nos promotores de seus genes alvos. Esses fatores se conectam em regiões distantes que contenham múltiplos sítios de ligação e que permitem sua ligação e interação com o complexo de transcrição. Essas regiões são conhecidas como regiões reguladoras.

1.2.3 Regiões reguladoras

Embora a região promotora de um gene seja responsável por ancorar a RNA Polimerase II e permitir o início de sua transcrição, muitas vezes a taxa de transcrição é muito baixa na ausência de uma região reguladora. Portanto, regiões reguladoras possuem um papel fundamental de facilitar e aumentar a velocidade de transcrição de um certo gene em específico.

As regiões reguladoras possuem, tipicamente, múltiplos sítios de ancoragem para diferentes fatores de transcrição ao longo de sua extensão. A ligação entre fatores de transcrição e regiões reguladoras ocorre através de motivos (*motifs*) – pequenas sequências definidas de nucleotídeos. Essas sequências funcionam como sítios de ancoragem e são reconhecidas pelos fatores de transcrição, permitindo a ligação destas proteínas a esse trecho específico na fita de DNA. Esses motivos podem ser encontrados em inúmeras regiões reguladoras do genoma, possuindo um alto grau de conservação. Essa característica possibilita que fatores de transcrição possam exercer controle sobre diversos genes de uma só vez (D'HAESELEER, 2006). Por sua vez, outras proteínas co-ativadoras ou co-repressoras podem interagir com fatores de transcrição conectados em cada região, formando um complexo proteico com um grande potencial de controle da transcrição gênica (SHLYUEVA; STAMPFEL; STARK, 2014).

Uma vez que esses diferentes fatores de ligação e suas proteínas acessório se ligam na região reguladora, há a formação de uma alça de DNA que permitirá a interação dessa região com o seu gene alvo, o qual pode estar há centenas de pares de bases acima ou abaixo de sua localização no genoma, ou ainda em outro cromossomo. Dessa maneira, a região reguladora irá carregar os fatores de transcrição e permitir o seu contato com o complexo de transcrição, de modo a modular a taxa de transcrição do gene alvo (BANERJI; RUSCONI; SCHAFFNER, 1981; TOLHUIS *et al.*, 2002).

A predição de áreas reguladoras é importante para o estudo da regulação gênica. Isso pode ser realizado através da investigação de sítios

de ligação para fatores de transcrição ao longo do genoma. Experimentos de sequenciamento de imunoprecipitação de cromatina (*Chromatin immunoprecipitation sequencing* - ChIP-Seq) são frequentemente utilizados para essa abordagem. Eles tornam possível a localização de fatores de transcrição ligados ao genoma através do uso de anticorpos, e o posterior sequenciamento dos nucleotídeos ancorados a eles. Consequentemente, os motivos de DNA associados a esses fatores de transcrição são revelados, assim como suas posições na sequência genética (JOHNSON *et al.*, 2007). Outra característica importante de áreas reguladoras, e que permite sua detecção, é o seu nível de acessibilidade do DNA. Devido ao fato dessas regiões necessitarem de ligação com diferentes proteínas, elas precisam estar constantemente acessíveis, isto é, não estarem compactadas. Portanto, um estudo do nível de compactação do genoma pode permitir a detecção de possíveis áreas reguladoras, auxiliando em seu estudo (SHLYUEVA; STAMPFEL; STARK, 2014).

1.2.4 Modificações de histonas

Os fatores epigenéticos possuem uma importante influência na ligação de fatores de transcrição com regiões reguladoras no genoma. A cromatina – a forma em que o DNA é compactado dentro da célula – é constituída de DNA envolto ao redor de proteínas, formando o nucleossomo. Essas proteínas, chamadas de histonas, possuem uma cauda de aminoácidos capaz de receber modificações pós-traducionais, como acetilações, metilações, fosforilações e ubiquitinações. Essas modificações são inseridas ou removidas dinamicamente por diferentes tipos de enzimas expressas pela célula (STRAHL; ALLIS, 2000). A histona H3, em especial, é conhecida por receber essas modificações químicas e causar mudanças no nível de compactação do DNA. Isso ocorre devido às alterações na carga elétrica da proteína e, por consequência, em seu nível de atração sob a cadeia nucleotídica (KOUZARIDES, 2007).

Dentre as modificações de cauda de histona conhecidas, a metilação na lisina 4 (H3K4me1), a tripla metilação na lisina 4 (H3K4me3), a acetilação na lisina 9 (H3K9ac) e na lisina 27 (H3K27ac) em caudas de histonas H3 são as mais estudadas. Por exemplo, as modificações do tipo H3K4me3 parecem estar envolvidas com áreas genômicas altamente transcritas (BERNSTEIN *et al.*, 2002). Por outro lado, alterações do tipo H3K4me1 (HEINTZMAN *et al.*, 2007), H3K9ac (KARMODIYA *et al.*, 2012) e H3K27ac (CREYGHTON *et al.*, 2010) são ligadas a regiões reguladoras ativas. Dessa maneira, é evidente que

modificações em caudas de histonas têm o potencial de influenciar diversos processos biológicos, de modo a ativar e desativar genes, ou suas regiões de influência. Isso corre através do aumento ou diminuição da acessibilidade do DNA à ligação de fatores de transcrição e/ou ao complexo de transcrição (KOUZARIDES, 2007). O reconhecimento do local dessas modificações e a sua característica química permite, então, a detecção de genes e de possíveis áreas reguladoras que estejam funcionalmente ativos em um determinado estado celular (HEINTZMAN *et al.*, 2007).

Uma outra maneira de detecção da acessibilidade do DNA é através da mensuração de áreas hipersensíveis à enzima DNase I. Essa enzima é capaz de digerir nucleotídeos que não estão compactados em nucleossomos, e que, portanto, estão potencialmente acessíveis à ligação do complexo de transcrição da RNA polimerase II. Por outro lado, trechos de DNA altamente envoltos ao redor de histonas não estão fisicamente suscetíveis à ação enzimática da DNase I ou à atividade de qualquer outra proteína, representando áreas inativas do genoma nuclear. Dessa maneira, quanto maior a sensibilidade à digestão da enzima DNase I, maior é a chance de a área em questão estar ativa geneticamente (GROSS; GARRARD, 1988). Ainda, regiões reguladoras geneticamente ativas são conhecidas por apresentarem ausência de nucleossomos, e, conseqüentemente, por terem uma grande sensibilidade à enzima DNase I (FELSENFELD *et al.*, 1996; TOLHUIS *et al.*, 2002).

1.3 O CONTROLE DA EXPRESSÃO DE *SOX2*

Ainda que o papel dos fatores de transcrição tenha sido intensamente estudado em diversos momentos da diferenciação celular, os mecanismos por trás do controle de sua expressão ainda não são completamente compreendidos. O fator de transcrição Sox2, em especial, possui um interessante perfil de expressão. Seu gene é altamente transcrito em células-tronco embrionárias. Porém, essa expressão é diminuída durante a diferenciação do embrioblasto nos tecidos embrionários da endoderme e mesoderme (progenitores da mesendoderme), enquanto continua ativo em tecidos da neuroectoderme (THOMSON *et al.*, 2011). Ao mesmo tempo, a superexpressão de *Sox2* em células-tronco embrionárias induz a sua diferenciação em neuroectoderme (KOPP *et al.*, 2008). Desse modo, parece que o controle minucioso da expressão gênica de *Sox2* é importante para conduzir a manutenção e o correto desenvolvimento dos tecidos embrionários. Isso é principalmente importante no desenvolvimento da neuroectoderme, que dará origem ao sistema nervoso central do embrião.

Além disso, Sox2 se mostra como um importante fator de transcrição na manutenção da pluripotencialidade de células-tronco embrionárias, permitindo a sustentação de culturas celulares *in vitro*. Por exemplo, a sua remoção completa da célula resulta na diferenciação de células-tronco embrionárias em células do trofoblasto (MASUI *et al.*, 2007). Ainda, seu papel celular parece ser exercido através da interação com outros fatores de transcrição – incluindo Oct4 e Nanog – demonstrando a existência de uma complexa rede de controle celular (MASUI *et al.*, 2007). Dessa maneira, a expressão do fator de transcrição Sox2 é um importante alvo de estudos para o melhor entendimento da rede de controle que governa a pluripotência de células-tronco.

Zhou *et al.* (2014) sugeriram que a expressão de Sox2 é controlada por diferentes regiões reguladoras. Uma possível região de controle principal foi denominada SCR (*Sox2 Control Region*). Ela abrange duas regiões reguladoras menores, a SRR (*Sox2 Regulatory Region*) 107 e a SRR111, localizadas a 107 e 111 mil pares de bases abaixo do promotor do gene, respectivamente (ZHOU *et al.*, 2014). Outras possíveis regiões de influência são a SRR1, SRR2 (TOMIOKA *et al.*, 2002) e SRR18 (ZHOU *et al.*, 2014), localizadas a 4 mil pares de bases acima, 4 mil pares de bases abaixo e 18 mil pares de bases abaixo do promotor do gene, respectivamente. Uma forma eficiente de atestar qual região reguladora é, de fato, crucial para a expressão de Sox2, seria através da geração de clones celulares contendo deleções gênicas nas principais regiões reguladoras desse gene. Dessa maneira, as consequências originadas pela remoção de cada região reguladora na expressão do gene Sox2 poderiam ser estudadas. Recentemente criado, o sistema de deleção CRISPR/Cas9 mostra-se como uma excelente ferramenta para esse tipo de abordagem.

1.3.1 O sistema de deleção CRISPR/Cas9

Inicialmente descoberto como um sistema de defesa imune adaptativo em células procariotas, o sistema CRISPR/Cas9 foi recentemente adotado como um interessante sistema de edição gênica em células eucariotas, permitindo a deleção de regiões genômicas inteiras facilmente.

Em bactérias, quando um material viral invasor é identificado no citoplasma da célula, ele é rapidamente clivado em trechos de cerca de 20 a 50 pares de bases e adicionado ao *loci* CRISPR no genoma bacteriano. Esse *loci* contém áreas espaçadoras, onde o DNA viral será incorporado e circundado por trechos repetitivos de DNA bacteriano. No final de cada sequência viral, há a adição de uma sequência específica chamada de PAM, de modo a sinalizar que esta é uma sequência adquirida. Isso

permite a criação de “vacinas” baseadas no registro de todo material genético previamente encontrado (BARRANGOU *et al.*, 2007; WIEDENHEFT; STERNBERG; DOUDNA, 2012). Quando uma nova infecção viral ocorre, todo o *loci* do CRISPR é expresso em um longo transcrito primário, o qual é prontamente cortado e processado em pequenos trechos de RNA contendo as sequências de DNA viral adquiridas de infecções prévias (crRNA). Esse crRNA irá, então, se associar à enzima Cas9, formando um complexo de clivagem nuclear CRISPR-Cas9. Caso a sequência viral presente no crRNA reconheça e se alinhe complementarmente a uma sequência invasora no citoplasma, o complexo irá realizar a quebra de dupla fita da molécula de DNA detectada (BROUNS *et al.*, 2008). Essa clivagem somente ocorre caso a sequência PAM esteja presente no alvo, de modo a evitar que o material da própria célula seja afetado (MOJICA *et al.*, 2009).

Esse sistema foi transferido para células eucarióticas de modo a permitir a edição precisa do seu material genético nuclear. Para realizar esse procedimento, três plasmídeos devem ser inseridos na célula. Um deles irá conter a enzima Cas9, enquanto os outros dois irão conter as sequências alvo complementares no genoma. Esses alvos devem ser escolhidos de modo a serem especificamente únicos no genoma e circundarem a área de deleção desejada. As sequências desses alvos deverão ser, então, inseridas dentro de um RNA guia (gRNA). O gRNA é um plasmídeo similar ao crDNA bacteriano, contendo as sequências acessórias necessárias para a correta sinalização da enzima Cas9. Dessa maneira, a enzima Cas9 irá detectar as sequências complementares e irá realizar uma quebra de dupla fita em dois pontos específicos no genoma – os alvos – removendo completamente o trecho presente entre esses locais. Os dois pontos de clivagem serão, então, unidos por recombinação não-homóloga e a região permanecerá deletada mesmo em futuras replicações celulares (SANDER; JOUNG, 2014). Desse modo, o sistema CRISPR/Cas9 permite a criação de linhagens celulares contendo deleções gênicas em diferentes áreas de interesse, possibilitando o estudo dos efeitos que essas deleções possam causar no fenótipo celular.

1.3.2 As regiões reguladoras de *Sox2*

Através do uso do sistema CRISPR/Cas9, Zhou *et al.* (2014) constataram que a remoção da região SCR em células-tronco embrionárias de camundongo resulta na completa ausência de expressão de *Sox2*, enquanto que deleções das regiões SRR1, SRR2 e SRR18 não demonstraram o mesmo efeito. Além disso, a perda da região SCR resultou numa remodelação do perfil de expressão gênica de células-

tronco embrionárias, sua alteração morfológica e perda da capacidade de geração de tecidos da ectoderme e neuroectoderme em corpos embrióides (ZHOU *et al.*, 2014). Esses resultados parecem indicar que a região SCR exerce um controle direto sobre a expressão de *Sox2* em células-tronco embrionárias.

Por outro lado, Ferri *et al.* (2004) demonstraram que a deleção da região SRR1 em camundongos *knock-out* resulta em más-formações cerebrais, embora células contendo essa deleção puderam formar e manter uma linhagem celular estável em cultura. Isso parece indicar que a região SRR1 é importante durante o controle do desenvolvimento cerebral, mas não é essencial para a manutenção da pluripotencialidade de células-tronco embrionárias (FERRI *et al.*, 2004). De modo similar, é possível que as regiões SRR2 e SRR18 possuam algum papel na regulação da expressão de *Sox2*, apesar de não serem completamente necessárias para manter sua expressão e a consequente conservação da pluripotencialidade de células-tronco embrionárias.

1.4 O PERFIL DE EXPRESSÃO GÊNICA DURANTE A TRANSIÇÃO PARA O EPIBLASTO

Ainda não está claro se a região SCR continua exercendo grande controle sobre a expressão de *Sox2* durante a diferenciação de células-tronco embrionárias, ou se outras regiões reguladoras assumem um maior papel nessa função.

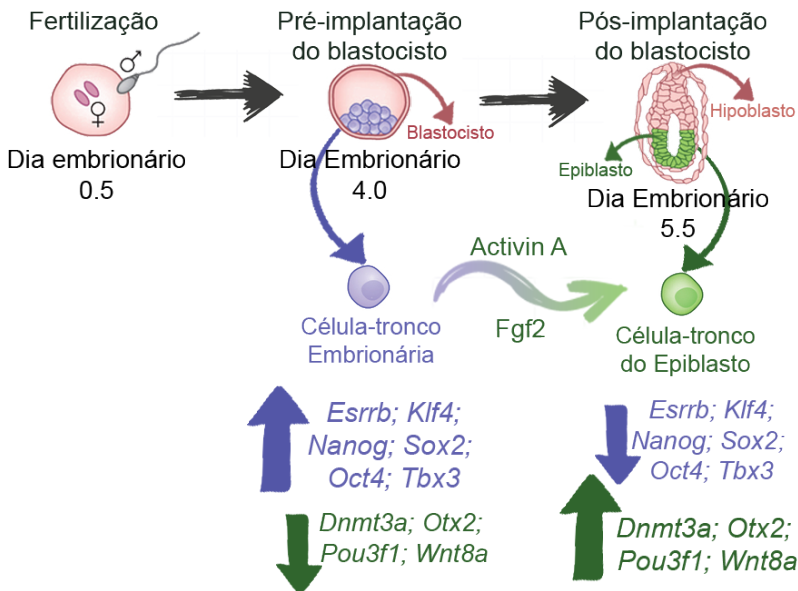
Buecker *et al.* (2014) conduziram a transição de células-tronco embrionárias para células-tronco induzidas a epiblasto utilizando os fatores de crescimento Activin A e Fgf2, essenciais para a manutenção de células do epiblasto (BRONS *et al.*, 2007). Após 48 horas de tratamento, os pesquisadores notaram uma considerável mudança na expressão gênica de diversos fatores de transcrição e fatores celulares relacionados ao desenvolvimento do blastocisto. Dentre eles, *Dnmt3a*, *Otx2*, e *Pou3f1* – fatores de transcrição associados com a pós-implantação do blastocisto – e *Wnt8a*, fator solúvel secretado pela célula, sofreram um aumento de expressão conforme o processo de diferenciação ocorreu. Por outro lado, *Esrrb*, *Klf4* e *Tbx3* – fatores de transcrição associados com a pré-implantação do blastocisto – sofreram uma redução em sua expressão mediante a diferenciação de células-tronco embrionárias. *Nanog*, *Sox2* e *Oct4* – fatores de transcrição relacionados à pluripotencialidade celular – também tiveram uma leve redução em sua expressão durante essa transição de estado celular (BUECKER *et al.*, 2014).

Essas variações são condizentes com o perfil de expressão gênica encontrado em células-tronco retiradas do epiblasto (BUECKER *et al.*,

2014). Por exemplo, o fator de transcrição Dnmt3a parece ser responsável pela metilação do DNA durante o desenvolvimento embrionário, silenciando genes conforme a diferenciação celular avança (OKANO *et al.*, 1999). Já Esrrb, Klf4, Nanog, Sox2 e Oct4 parecem atuar juntos e serem cruciais para a manutenção da pluripotencialidade de células-tronco embrionárias, pois a remoção de qualquer um desses componentes resulta na diferenciação celular (BOYER *et al.*, 2005; JIANG *et al.*, 2008; ZHANG *et al.*, 2008). Além disso, Tbx3 também auxilia no bloqueio da diferenciação de células-tronco embrionárias aos três folhetos embrionários (IVANOVA *et al.*, 2006). Por outro lado, Otx2 parece ser fundamental para a correta diferenciação de células-tronco embrionárias em células-tronco do epiblasto em conjunto com a sinalização de Fgf2 (ACAMPORA; GIOVANNANTONIO; SIMEONE, 2012). Ainda, Pou3f1 aparenta ser responsável por conduzir a diferenciação de células-tronco em progenitores neurais, sendo altamente expresso em tecidos neurais em desenvolvimento (ZHU *et al.*, 2014). Por fim, a expressão do gene *Wnt8a*, pertencente à família de moléculas de sinalização Wnt, contribui para a formação da mesoderme e da linha primitiva a partir do epiblasto durante a gastrulação (LIU *et al.*, 1999; LINDSLEY *et al.*, 2006). Portanto, o nível de expressão desses genes parece ser um bom indicador da mudança de estado celular entre células-tronco embrionárias e células-tronco induzidas a epiblasto. Um resumo do desenvolvimento embrionário, da diferenciação celular, assim como a variação de fatores celulares envolvidos nesse processo, pode ser visto na Figura 1.

Conforme essa metodologia, células-tronco embrionárias que contenham deleções gênicas em regiões reguladoras de *Sox2* podem ser diferenciadas em células-tronco induzidas a epiblasto mediante o uso de Activin A e Fgf2. Desse modo, será possível verificar se a região SCR continua fundamental para a expressão de *Sox2* mesmo durante a diferenciação celular, ou se as regiões SRR1, SRR2 e SRR18 assumem um maior papel nesse controle. Além disso, os genes citados anteriormente podem ser analisados de modo a confirmar a ocorrência desse processo de diferenciação.

Figura 1: Desenvolvimento embrionário inicial de camundongos. Células-tronco embrionárias são extraídas da massa interna de células durante a fase de pré-implantação do blastocisto e possuem maior expressão gênica dos fatores de transcrição *Esrrb*, *Klf4*, *Nanog*, *Sox2*, *Oct4* e *Tbx3*. Já células-tronco do epiblasto são extraídas do epiblasto, formado após a implantação do blastocisto. Nesse estado, há a maior expressão gênica dos fatores de transcrição *Dnmt3a*, *Otx2*, *Pou3f1* e do fator celular *Wnt8a*. A diferenciação de células-tronco embrionárias em células-tronco do epiblasto é possível através da exposição aos fatores de crescimento Activin A e Fgf2 (BUECKER *et al.*, 2014). Imagem modificada de Bock e Wutz (2013).



Fonte: O autor (2015).

1.5 OS FATORES DE TRANSCRIÇÃO DENTRO DE REGIÕES REGULADORAS

Demonstrar que a região SCR é fundamental para o controle da expressão de *Sox2* em células-tronco induzidas a epiblasto é tão importante quanto conhecer a sua real composição. Se a região SCR é, de fato, uma região reguladora, ela deve incluir motivos conservados de ligação para fatores de transcrição, os quais irão, por sua vez, promover a expressão de *Sox2*. De forma análoga, a região reguladora do gene de beta-globina, por exemplo, possui diversos motivos de DNA que permitem a ancoragem de diversas proteínas, criando um sítio de

interação proteica que resulta na correta expressão do gene (LEVINGS; BUNGERT, 2002).

Dessa maneira, é imprescindível investigar quais são os fatores de transcrição atuantes na região SCR e onde esses fatores de transcrição se ligam dentro dessa região reguladora, bem como investigar a possível interação entre eles. Isso permite elucidar a real função que a região SCR promove sobre a expressão do gene *Sox2 in vivo*.

1.6 JUSTIFICATIVA

A compreensão dos mecanismos de controle da expressão dos genes que mantêm a pluripotencialidade de células-tronco embrionárias se mostra importante devido ao grande potencial que essas células representam dentro da medicina regenerativa e no desenvolvimento embrionário. Nesse contexto, o melhor entendimento de como a célula regula a expressão de *Sox2* pode permitir a indução de sua expressão em células já diferenciadas e possibilitar um avanço em pesquisas e tratamentos que envolvam a utilização de células-tronco. Uma outra possibilidade é proporcionar um melhor controle da diferenciação de células-tronco embrionárias, possibilitando o seu direcionamento para a formação de tecidos específicos do corpo.

1.7 HIPÓTESES

No presente trabalho, a influência das regiões reguladoras SRR1, SRR2, SRR18 e SCR sob a expressão do fator de transcrição *Sox2* foi investigada em células-tronco embrionárias e células-tronco induzidas a epiblasto. Diferentes hipóteses foram testadas. Primeiramente, a região SCR é a região reguladora mais importante para o controle da expressão da proteína Sox2 em células-tronco embrionárias. Em seguida, a transição entre células-tronco embrionárias e células-tronco induzidas a epiblasto através do uso de Activin A e Fgf2 reflete na mudança do padrão de expressão dos genes *Dnmt3a*, *Esrrb*, *Klf4*, *Nanog*, *Oct4*, *Pou3f1*, *Otx2*, *Tbx3* e *Wnt8a*. Posteriormente, a região SCR continua sendo a região reguladora mais importante para a expressão de *Sox2* em células-tronco induzidas a epiblasto. Por ser tão importante, a região SCR exibe características epigenéticas condizentes com regiões reguladoras em comum, como acetilações e metilações em caudas de histonas H3 e grande acessibilidade do DNA. Ainda, a região SCR interage com diversos fatores de transcrição para promover a expressão de *Sox2*. Com isso, a região SCR se revela como a principal região reguladora para a expressão de *Sox2* em células-tronco embrionárias e células-tronco induzidas a epiblasto de camundongo.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Investigar a regulação da expressão do gene *Sox2* em células-tronco embrionárias, antes e após sua indução a epiblasto de camundongo.

2.2 OBJETIVOS ESPECÍFICOS

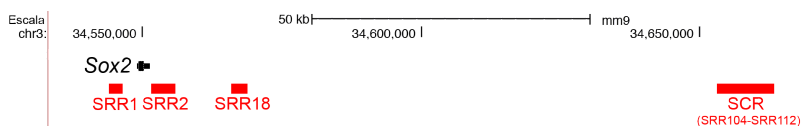
1. Verificar os efeitos da deleção das regiões reguladoras SRR1, SRR2, SRR18 e SCR sobre a expressão da proteína *Sox2*, em células-tronco embrionárias.
2. Conduzir a diferenciação de células-tronco embrionárias para células-tronco induzidas a epiblasto, em modelo murino.
3. Analisar a variação de genes indicadores do estado celular, *Dnmt3a*, *Esrrb*, *Klf4*, *Nanog*, *Oct4*, *Pou3f1*, *Otx2*, *Tbx3* e *Wnt8a*, durante a diferenciação ao epiblasto.
4. Comparar os efeitos causados pela deleção heterozigótica da região SCR e das regiões SRR1, SRR2 e SRR18 na expressão de *Sox2* durante a transição entre células-tronco embrionárias e células-tronco induzidas a epiblasto.
5. Estudar as características epigenéticas de regiões reguladoras que controlam a expressão de *Sox2* em células-tronco embrionárias.
6. Investigar possíveis fatores de transcrição que se liguem às regiões reguladoras que apresentem influência na expressão de *Sox2*, em células-tronco embrionárias.

3 MATERIAL E MÉTODOS

Os procedimentos realizados foram divididos em três blocos experimentais. No primeiro bloco, células-tronco embrionárias foram cultivadas para a extração de proteínas, seguido da detecção dos níveis de *Sox2* comparando-o com o controle *Gapdh*, o qual apresenta expressão estável e constante. No segundo bloco, novas células-tronco foram cultivadas para a avaliação do RNA mensageiro e determinação dos níveis de expressão de fatores de transcrição, incluindo *Sox2*. O terceiro bloco, por sua vez, envolve experimentos *in silico* para a determinação dos mecanismos de controle gênico da expressão do gene *Sox2* e de fatores de transcrição que possam influenciar esse processo.

A figura 2 ilustra a posição relativa das áreas ocupadas pelas regiões reguladoras SRR1, SRR2, SRR18 e SCR (SRR104 a SRR112) no genoma do camundongo. Em ambos os blocos experimentais 1 e 2, células-tronco embrionárias tiveram deleções em uma ou mais dessas áreas, conforme indicado. Cada procedimento de deleção produziu diferentes clones com pequenas variações na posição de deleção, devido às limitações do sistema de deleção CRISPR/Cas9 (ZHOU *et al.*, 2014).

Figura 2: Mapa genômico da posição das regiões reguladoras SRR1, SRR2, SRR18 e SCR (SRR104 a SRR112) em relação ao gene alvo *Sox2* (ZHOU *et al.*, 2014) no cromossomo 3 de camundongos. A escala de 50.000 pares de bases (kb) é representada abaixo. Essa ilustração foi gerada através do mapa genômico da UCSC, utilizando-se a compilação do genoma de camundongo mm9 (KENT *et al.*, 2002).



Fonte: KENT *et al.* (2002)

3.1 PRIMEIRO BLOCO EXPERIMENTAL

3.1.1 O cultivo de células-tronco embrionárias

Diferentes clones celulares foram previamente produzidos utilizando-se o sistema de deleção gênica CRISPR/Cas9 no laboratório de estudos de controle gênico da Doutora Jennifer A. Mitchell na Universidade de Toronto, Canadá. Essas células eram provenientes do cruzamento de camundongos *M. musculus*¹²⁹ x *M. castaneus* e possuíam como alvo de deleção as regiões candidatas para o controle da expressão de *Sox2*, incluindo as regiões SRR1, SRR2, SRR18, SCR e o próprio gene *Sox2* (ZHOU *et al.*, 2014). Embora a maioria das deleções afetou ambos

os alelos (*129* e *Castaneus*), alguns clones possuíam deleções heterozigóticas, ou seja, apenas no alelo *Castaneus*. Além disso, clones com deleções em mais de uma região candidata simultaneamente também foram utilizados para análise. Uma lista da nomenclatura adotada para nomear os grupos de deleções gênicas pode ser apurada no Quadro 1. A sequência nucleotídica das deleções realizadas em cada grupo de células pode ser verificada no Anexo A.

Todos os procedimentos adotados para a extração de células-tronco proveniente de camundongos tiveram a aprovação emitida pelo Comitê de Cuidado Animal da Universidade de Toronto (*University of Toronto Animal Care Committee - UACC*) e do Comitê de Cuidado Animal para o uso em pesquisas de Biociências (*Bioscience Local Animal Care Committee – LACC*) sob o número de licença 527-M22-1.

Quadro 1: Relação da nomenclatura de deleções realizadas em células-tronco embrionárias através do sistema CRISPR/Cas9, clones celulares gerados, alelos deletados e regiões reguladoras afetadas por cada deleção.

Nome da deleção	Clones existentes e alelos removidos	Regiões reguladoras afetadas
Δ SCR	Clone 1 – Alelo <i>129</i> e <i>Castaneus</i>	SCR (SRR104 a SRR112)
Δ SRR1/2/18	Clone 39 – Alelo <i>Castaneus</i>	SRR1 + SRR2 + SRR18
Δ SRR1 + Δ SRR18	Clone 25 e Clone 39 – Alelo <i>129</i> e <i>Castaneus</i> .	SRR1 + SRR18
Δ SRR2	Clone 17, Clone 30, Clone 54, Clone 62 e Clone 91 – Alelo <i>129</i> e <i>Castaneus</i>	SRR2
Δ SRR2-SRR104	Clone 18 e Clone 21 – Alelo <i>129</i> e <i>Castaneus</i>	SRR2 a SRR104
Δ SRR18	Clone 1, Clone 4, Clone 30, Clone 36 e Clone 37 – Alelo <i>129</i> e <i>Castaneus</i>	SRR18
Δ SRR18-SRR104	Clone 10, 21 e 39 – Alelo <i>Castaneus</i> / Clone 70 – Alelo <i>129</i> e <i>Castaneus</i>	SRR18 a SRR104
Δ SOX2	Clone 36 – Alelo <i>Castaneus</i>	Gene <i>Sox2</i>
F1	-	Nenhuma

Fonte: O autor (2015).

As células-tronco embrionárias portadoras das deleções gênicas citadas foram previamente produzidas, replicadas e congeladas (ZHOU *et al.*, 2014). Para realizar os experimentos, essas células foram descongeladas do estoque de nitrogênio líquido e cultivados em placas de cultura celular de 10cm² previamente tratadas e cobertas com 0.1% de gelatina. Foi utilizado meio de cultura ES (Gibco® DMEM, 15% Gibco® Soro Fetal Bovino, 0.1mM Thermo Fisher Scientific™ MEM aminoácidos não-essenciais, 1mM Gibco® Piruvato de Sódio, 2mM Thermo Fisher Scientific™ GlutaMAX®, 0.1mM 2-mercaptoethanol, 3 µM CHIR99021 [Inibidor GSK3β, Biovision™], 1 µM PD0325901 [Inibidor MEK, Invitrogen™] e 1000 U/ml LIF [Fator Inibidor de Leucemia, Sigma-Aldrich™]) durante todo o procedimento (ZHOU *et al.*, 2014). As células foram mantidas em incubadora celular à 37°C com 5% de CO₂. Cada placa de cultivo celular teve o seu meio de cultura trocado até a densidade celular atingir aproximadamente 10⁶ células. A contagem de células foi realizada utilizando-se contador celular automático (ORFLO® MOXI Z *Mini Automated Cell Counter*).

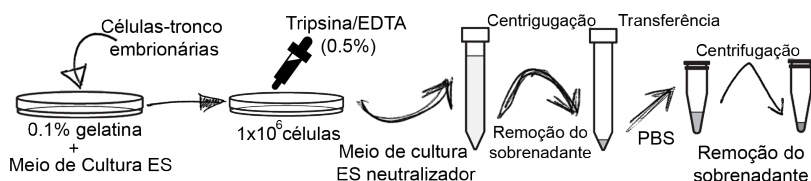
Com a densidade celular desejada alcançada, as células foram tratadas com solução de tripsina/EDTA (0.5% tripsina) e incubadas por 5 minutos. Posteriormente, meio de cultura com neutralizador de tripsina (Gibco® DMEM, 15% Gibco® Soro Fetal Bovino, 1mM Gibco® Piruvato de Sódio e 2mM Thermo Fisher Scientific™ GlutaMAX®) foi utilizado para transferir as células para tubos de 15ml. Esses tubos foram centrifugados à 1400 rpm por 4 minutos. Após esse procedimento, o meio de cultura sobrenadante foi removido e 1ml de PBS (Gibco® 10X Tampão Fosfato-Salino) foi utilizado para ressuspender e transferir o conteúdo celular para tubos de 1,5ml. Esses tubos foram novamente centrifugados à 2.000 rpm e o sobrenadante removido. O aglomerado final de células foi então congelado a -80°C até a próxima etapa de experimentos. Um resumo desse procedimento pode ser visto na Figura 3.

3.1.2 Extração e cálculo de concentração proteica

Após o cultivo e congelamento de todos os clones celulares, cada um deles foi retirado do congelador e diretamente tratado com uma solução de extração de proteínas 1X RIPA (Thermo Fisher Scientific™) com adição de inibidores de proteases (Roche™ *cOmplete Protease Inhibitor Cocktail* + Thermo Fisher Scientific™ *PMSF Protease Inhibitor*). Cada tubo foi centrifugado à 14.000 rpm por 15 minutos. O sobrenadante, contendo o extrato proteico, foi transferido para um novo tubo.

A concentração aproximada de proteínas resultante da extração foi calculada utilizando o sistema de reagentes BCA (Thermo Fisher Scientific™ Pierce® *BCA Protein Assay*). Nele, uma amostra de cada extrato proteico foi adicionada aos reagentes provenientes do kit e encubados à 37°C por 30 minutos. Após o período de incubação, a concentração proteica foi determinada em espectrofotômetro através da leitura de absorbância em 562nm.

Figura 3: Síntese dos procedimentos utilizado para a preparação das células-tronco embrionárias utilizadas durante o primeiro bloco experimental. Imagem modificada de Zhang e Kuhn (2013).



Fonte: O autor (2015).

3.1.3 Detecção da proteína Sox2 em relação ao controle Gapdh

O kit de imunodeteção proteica – *Western Blot* (Thermo Fisher Scientific™ Bolt® *Mini Gel*) – foi utilizado para a eletroforese em gel SDS pré-fabricado das amostras de proteína. Com a concentração proteica de cada solução conhecida, foi possível padronizar as amostras baseando-se na menor concentração dentre elas. Dessa forma, foi realizada a diluição de cada amostra em água destilada de forma a obter um volume final de 10ul. Essa diluição foi incubada à 75°C para desnaturação proteica e, em seguida, adicionada a cada poço no gel.

Devido à necessidade da obtenção de uma reação imunológica controle (Gapdh – 335 aminoácidos, massa molecular 36053 Da) para comparação com a detecção da proteína alvo (Sox2 – 317 aminoácidos, massa molecular 34310 Da), foi feita uma duplicata para cada amostra. Isto é, cada amostra foi adicionada em dois géis diferentes, na mesma posição e ordem. Dessa forma, foram realizadas 8 eletroforeses, sendo 4 delas duplicatas. Todas as oito passaram pelo mesmo procedimento de corrida por 1 hora à 100V, com os outros parâmetros elétricos estáveis.

Em seguida, os géis foram removidos e lavados em uma solução de metanol, tris/glicina e água. Oito membranas de PVDF (Thermo Fisher Scientific™) foram carregadas aplicando-se metanol (100%). Cada

membrana foi posicionada junto a cada gel para a transferência do material proteico. Visando evitar a formação de bolhas de ar entre o gel e a membrana de PVDF, esse procedimento foi realizado em recipiente específico, onde os materiais ficaram imersos em tampão de transferência (Thermo Fisher Scientific™). O processo de transferência foi realizado durante 18 horas a 25V.

Após o término da transferência, as membranas de PVDF foram transferidas para uma solução de bloqueio de proteínas (Tween20, TBS e leite desnatado em pó), sob agitação leve, por 45 minutos.

Posteriormente, cada membrana foi lavada com uma solução de TTBS (1x TBS + 0.05% Tween20) durante 10 minutos por três vezes consecutivas. Na sequência, anticorpos primários para Sox2 (ABCAM™, ab75485, provenientes de camundongo) foram adicionados junto a TTBS (diluição 1:500) e colocados sob 4 membranas, sendo que suas duplicatas passaram pelo mesmo procedimento, mas foram banhadas utilizando anticorpos primários para Gapdh (ABCAM™, ab8245, provenientes de camundongo), com a mesma diluição. Todas as membranas permaneceram expostas aos seus respectivos anticorpos durante 18 horas à 4°C.

Em seguida, cinco novas sequências de lavagem de 5 minutos com TTBS foram realizadas. Cada membrana foi, então, exposta aos anticorpos secundários – conjugados a peroxidase – reativos para proteínas de camundongo (ABCAM™, ab150113, provenientes de cabra) em diluição com TTBS de 1:1000. Após 45 minutos de incubação, o reagente de iluminação (Biorad™ Clarity® Western ECL *Blotting Substrate*) foi aplicado sob as membranas e levadas ao fotodocumentador (Bio-Rad™ ChemiDoc® *Imaging System*). Cada membrana foi fotografada e as duplicatas foram comparadas através do *software* Bio-Rad™ Image Lab®, onde foi possível a visualização das bandas e da formação de imagens em três dimensões de cada amostra proteica. A construção das imagens em três dimensões reflete a intensidade das bandas de cada amostra na membrana. Isso permitiu determinar os níveis da proteína Sox2, comparados com a presença da proteína Gapdh em cada clone celular. Essa proteína controle permite a comparação dos efeitos das deleções de cada região reguladora sobre a expressão da proteína Sox2 em cada extrato proteico.

O *software* Photoshop foi utilizado para a edição das imagens de imunodeteção e para o cálculo aproximado do volume proteico de cada amostra celular em relação ao controle F1 e à proteína Gapdh. Esse volume corresponde à área do gráfico em três dimensões gerado previamente, dividido pela área do gráfico de células controle F1.

3.2 SEGUNDO BLOCO EXPERIMENTAL

3.2.1 O cultivo de células-tronco embrionárias

Células-tronco embrionárias da geração F1 do cruzamento de camundongos *M. musculus*¹²⁹ x *M. castaneus* e contendo deleções heterozigóticas em um de seus alelos (*Castaneus*) nas regiões SCR (Grupo 1 - Δ SCR^{Cast} - Clone 11), nas regiões SRR1, SRR2 e SRR18 simultaneamente (Grupo 2 - Δ SRR1/2/18^{Cast} - Clone 39), assim como células controle sem deleções (Grupo 3 – Clone Controle) foram obtidas dos estoques de nitrogênio líquido do laboratório da Doutora Jennifer A. Mitchell na Universidade de Toronto, Canadá. A utilização de células híbridas permite o cálculo da expressão de genes localizados em cada alelo separadamente, e como as deleções os afetam separadamente.

Todos os clones celulares foram descongelados e cultivados em placas de 10cm² previamente tratadas e cobertas com 0.1% de gelatina, utilizando-se o meio de cultura N2B27 (Gibco® DMEM/F12 GlutaMAXTM 1:1 Meio de Cultura Gibco® Neurobasal®, 1X Suplemento Gibco® N-2, 1X Suplemento Gibco® B-27®, 0.1 mM 2-mercaptoethanol, 1X Gibco® *Penicillin-Streptomycin*) complementado com os inibidores de diferenciação 3 μ M CHIR99021 (Inibidor GSK3 β , BiovisionTM) e 1 μ M PD0325901 (Inibidor MEK, InvitrogenTM) de modo a manter a pluripotência celular sem uma camada extra de células alimentadoras (*feed layer free*) (YING; SMITH, 2003; YING *et al.*, 2008; ZHOU *et al.*, 2014). As placas foram mantidas à 37°C em uma incubadora celular configurada à 5% de CO₂.

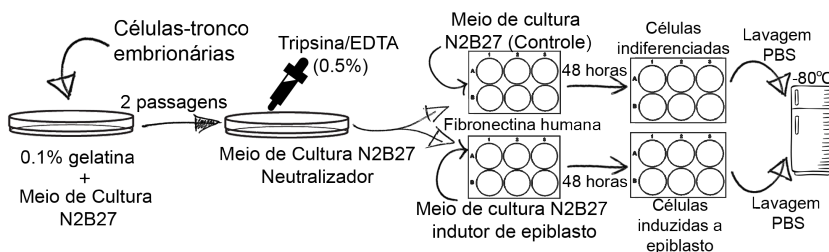
O meio de cultura foi trocado a cada dois dias, ao mesmo tempo em que as células foram passadas para novas placas. Esse procedimento, denominado “passagem”, constitui-se do uso de solução de tripsina/EDTA (0.5% tripsina) para a separação das células do fundo das placas seguido de incubação por 5 minutos à 37°C. Após o tempo decorrido, um novo meio de cultura N2B27 com neutralizador de tripsina (Meio de cultura N2B27 Modificado – Gibco® DMEM/F12 GlutaMAX 1:1 Meio de Cultura Gibco® Neurobasal®, 1% Thermo Fisher ScientificTM KnockOut® *Serum Replacement*, 1X Gibco® *Penicillin-Streptomycin*) foi adicionado de modo a permitir a transferência das células para novos tubos de 15ml. Esses tubos foram centrifugados à 1400 rpm por 4 minutos. O meio de cultura de neutralização foi removido e um novo meio de cultura N2B27 foi adicionado em cada tubo para permitir a ressuspensão das células. O número de células foi contado utilizando-se contador celular automático (ORFLO® MOXI Z Mini Automated Cell Counter), e cerca de 10⁵ células foram aplicadas em novas placas de cultura celular e guardadas no incubador.

3.2.2 Indução do estado de epiblasto

Após duas passagens, cada grupo de células foi dividido em dois subgrupos. Ambos subgrupos foram transferidos para placas de cultivo celular com 6 poços (*6 well plates*) previamente revestidos com 10µg/ml de fibronectina humana (Millipore® *human plasma fibronectin*). Um dos subgrupos continuou recebendo meio de cultura N2B27 (grupo controle), enquanto o outro recebeu um meio de cultura N2B27 modificado para indução ao estado de epiblasto (*Epiblast-like cell media inducer* – Gibco® DMEM/F12 GlutaMAX™ 1:1 Meio de Cultura Gibco® Neurobasal®, 1% Thermo Fisher Scientific® KnockOut™ Serum Replacement, 1X Gibco® Penicillin-Streptomycin, 12ng/ml PeproTech® recombinant murine Fgf2, 20ng/ml R&D Systems® Recombinant human/mouse/rat Activin A), como descrito na literatura (HAYASHI *et al.*, 2011; BUECKER *et al.*, 2014).

Cada subgrupo teve seu meio de cultura removido após 48 horas de diferenciação celular. Posteriormente, uma solução de PBS (Gibco® 10X Tampão Fosfato-Salino) à 4°C foi cuidadosamente aplicada sobre as placas de cultura para a lavagem e remoção de eventuais células mortas. Depois da remoção do PBS, as placas de cultura foram seladas e congeladas à -80°C para a conservação celular até o próximo passo experimental. Três diferentes réplicas biológicas foram geradas utilizando-se o mesmo protocolo. Um resumo do protocolo de cultivo celular pode ser visto na Figura 4.

Figura 4: Síntese dos procedimentos utilizados para a preparação das células-tronco embrionárias e células-tronco induzidas a epiblasto utilizadas durante o segundo bloco experimental. Imagem modificada de Zhang e Kuhn (2013).



Fonte: O autor (2015).

3.2.3 Extração de RNA

A extração do RNA celular total foi realizada utilizando-se o kit Qiagen® RNeasy Mini Kit. Nesse procedimento, as células são lisadas,

homogeneizadas, solidificadas em etanol e centrifugadas através da passagem por um filtro de captura de RNA. Posteriormente, o RNA resultante é eluído em 40µl de água ultrapurificada (Sigma-Aldrich® *Water*). Enzimas hidrolíticas de DNA (BioLabs® - DNase I) foram adicionadas de modo a eliminar o DNA residual da amostra, completando 50µl de solução. A solução foi incubada à 37°C por 10 minutos. Em seguida, a ação da enzima de DNA foi interrompida através da adição de EDTA e aquecimento à 75°C por 10 minutos. O RNA resultante desse processo foi quantificado através de espectrofotometria 260/280nm em equipamento NanoDrop® (Thermo Fisher Scientific).

Cerca de 500ng de RNA serviram de base para a transcrição reversa em DNA complementar utilizando-se o kit iScript™ (Bio-Rad®). Esse kit contém a enzima transcriptase reversa e um mix de reação, além de *primers* randômicos (Thermo Fisher Scientific) de modo a amplificar toda a extensão do transcriptoma. O volume total da reação foi de 20µl, a qual foi colocada no termociclador. O ciclo termal adotado foi de 5 minutos à 25°C, 30 minutos à 42°C e 5 minutos à 85°C, de modo a efetivar a reação conforme as especificações do fabricante. Após esse processo, as amostras foram armazenadas à 4°C até a próxima etapa experimental.

3.2.4 Desenho de primers

As sequências dos genes *Dnmt3a*, *Esrrb*, *Gapdh*, *Klf4*, *Nanog*, *Oct4*, *Pou3f1*, *Otx2*, *Sox2*, *Tbx3* e *Wnt8a* foram obtidas através do mapa genômico UCSC (KENT *et al.*, 2002). As sequências dos éxons de cada gene foram inseridas no *software* Primer3Plus (UNTERGASSER *et al.*, 2007) para a obtenção de *primers* candidatos. Esses *primers* foram sintetizados pela Life Technologies. Após sua entrega, cada *primer* foi solubilizado em concentrações de 100µM e armazenados a 4°C. Diluições adicionais de 10 µM foram criadas para utilização nos próximos experimentos através da adição de água ultrapurificada (Sigma-Aldrich® *Water*). Posteriormente, cada *primer* foi testado em amplificação de PCR quantitativo utilizando-se DNA genômico de células F1, de modo a exibir sua eficácia de amplificação. Somente *primers* que possuíam alta especificidade e curva de amplificação ideais foram selecionados. Todos *primers* amplificavam ambos os alelos *129* e *Cast*, com exceção do *primer* para o fator de transcrição *Sox2*. Nesse caso, dois *primers* distintos foram construídos, cada um capaz de amplificar um dos alelos. Desse modo, foi possível quantificar distintamente a expressão de *Sox2* proveniente de cada alelo. As sequências finais dos *primers* utilizados estão apresentados no Quadro 2.

Quadro 2: Sequência de *primers* utilizadas para os experimentos de PCR quantitativo de transcrição reversa (RT-qPCR) desenhados através do uso do Primer3Plus (UNTERGASSER *et al.*, 2007).

Gene	Sequência <i>Primer</i> Senso	Sequência <i>Primer</i> Anti-senso
<i>Dnmt3a</i>	5' cttccccgtccactacacag	5' aagaggtggcggatgacc
<i>Esrrb</i>	5' ctcattctgggcatcgtga	5' ctaccaggcgagagtgttcc
<i>Gapdh</i>	5' gcaccagcatccctagacc	5' cttctgtgcagtgccagggtg
<i>Klf4</i>	5' gaagacgaggatgaagctgac	5' tggacctagactttatctcttcc
<i>Nanog</i>	5' tcccaaacaaaagctctcaag	5' atctgctggaggctgaggta
<i>Oct4</i>	5' atgaggctacagggacacctt	5' gtgaagtgggggcttcata
<i>Otx2</i>	5' gaagttgagccagcatagcc	5' tctgacccttgtccacttc
<i>Pou3f1</i>	5' caagaagcgcacgtccat	5' agacggcttgggacacttg
<i>Sox2_129</i>	5' ggactctttttgggggact	5' cgctaactgtaccactagaacttt
<i>Sox2_Cast</i>	5' ggactctttttgggggact	5' cgctaactgtaccactagaactta
<i>Tbx3</i>	5' ggaacccgaagaagacgtaga	5' ctgtaccccgcttgtaaacc
<i>Wnt8a</i>	5' ggacaagcgccagctaag	5' cgttgcggttgtagtagtc

Fonte: O autor (2015).

3.2.5 Quantificação da expressão gênica

A quantificação da expressão gênica foi realizada através do uso de PCR quantitativo de transcrição reversa (*reverse transcription quantitative polymerase chain reaction – RT-qPCR*). Nesse procedimento, 1µl de amostra de DNA complementar foi adicionada aos *primers* específicos para cada fator de transcrição (*Forward* e *Reverse*, 1µl cada de solução à 10 µM), 5µl de SYBR® *Select Master Mix* e 3µl de Água Sigma® em placas contendo 384 poços. O equipamento utilizado para a leitura da expressão gênica foi o Bio-Rad® CFX384 Touch™ *Real-Time PCR Detection System*. O experimento foi desenhado de modo que cada placa de amplificação correspondia a um dos fatores de transcrição, contendo 3 réplicas técnicas para cada amostra. Desse modo, foi possível obter uma média final de expressão gênica de cada fator de transcrição baseado nos três grupos de células (F1, ΔSCR^{Cast} e ΔSRR1/2/18^{Cast}) e seus dois tratamentos (controle e induzido a epiblasto) provenientes de cada réplica biológica. Os valores de expressão gênica iniciais foram divididos pela expressão do gene controle *Gapdh* de modo

a normalizar os níveis de expressão entre diferentes grupos de células (BARBER *et al.*, 2005).

3.2.6 Análise estatística do perfil de expressão gênica

Os dados obtidos através dos experimentos de PCR quantitativo de transcrição reversa foram analisados utilizando o *software* GraphPad Prism 6.0 para Windows. O teste estatístico de análise de variâncias de uma via (*one-way ANOVA*) foi escolhido para verificar as diferenças entre as médias de expressão de cada gene obtidas de cada réplica biológica. O pós-teste (*post hoc*) de análise de médias de Tukey foi aplicado. Valores de p menores que 0,05 foram considerados significativos.

Gráficos contendo as médias de expressão de cada gene foram gerados utilizando o *software* GraphPad Prism 6.0 para Windows. Neles, cada um dos três grupos de células foi comparado entre seus respectivos estados de células-tronco embrionárias (controle) e células-tronco induzidas a epiblasto. Valores de p menores que 0,05 foram destacados em cada gráfico, quando existentes.

3.3 TERCEIRO BLOCO EXPERIMENTAL

3.3.1 O papel molecular da região SCR

Para confirmar a possível função reguladora da região SCR, modificações epigenéticas de histonas H3 presentes nessa região foram verificadas através de experimentos de sequenciamento de imunoprecipitação de cromatina (ChIP-Seq) realizados pelo consórcio ENCODE (MOUSE ENCODE CONSORTIUM *et al.*, 2012). Esses dados foram obtidos *in silico* através do mapa genômico da Universidade da Califórnia Santa Cruz, utilizando a versão mm9 do genoma do camundongo (UCSC Genome Browser) (KENT *et al.*, 2002). Dados que indicam as metilações na lisina 4 da cauda de histonas H3 (H3K4me1), trimetilações na lisina 4 da cauda de histonas H3 (H3K4me3), acetilações na lisina 9 da cauda de histonas H3 (H3K9ac) e acetilações na lisina 27 da cauda de histonas H3 (H3K27ac) foram utilizados (MOUSE ENCODE CONSORTIUM *et al.*, 2012).

O gene *Sox2* e as regiões reguladoras SRR1, SRR2, SRR18 e SCR foram plotados no mapa genômico UCSC em suas respectivas posições cromossômicas. Essas posições foram sinalizadas buscando a sequência de cada região reguladora, presentes no Anexo A, dentro do banco de dados do mapa genômico (KENT *et al.*, 2002). Em seguida, a intensidade de sinal obtida de cada experimento de sequenciamento de imunoprecipitação foi alinhada à essas regiões, permitindo comparar quais regiões do genoma possuem a maior quantidade de cada modificação de histona H3, sejam ela acetilações ou metilações.

O mesmo procedimento foi realizado em relação aos dados de hipersensibilidade à enzima DNase I, detecção de modificações de histona e experimentos de ligação de fatores de transcrição, novamente obtidos através do consórcio ENCODE (MOUSE ENCODE CONSORTIUM *et al.*, 2012). Os resultados foram alinhados às posições de cada região reguladora e ao gene *Sox2*. As imagens foram editadas no *software* Photoshop CS6. As cores originais foram mantidas, enquanto a região SCR foi destacada na imagem.

3.3.2 Os fatores de transcrição relacionados à região SCR

De modo a encontrar possíveis fatores de transcrição que atuam na região SCR, experimentos de sequenciamento de imunoprecipitação realizadas em células-tronco embrionárias por Chen *et al.* (2008) foram consultados *in silico* através do banco de dados CODEX (CHEN *et al.*, 2008; SANCHEZ-CASTILLO *et al.*, 2014).

Ao todo, dez fatores de transcrição – Ctf, E2f1, Esrrb, Klf4, Nanog, Oct4, Smad1, Sox2, Stat3 e Tfc2l1 – tiveram seus picos de atividade plotados no mapa genômico da Universidade da Califórnia Santa Cruz (UCSC Genome Browser) em relação às posições do gene *Sox2* e das regiões estimuladores SRR1, SRR2, SRR18 e SCR (KENT *et al.*, 2002). A versão mm10 do genoma do camundongo foi utilizada para fins de compatibilidade com os dados gerados pelos experimentos de imunoprecipitação (CHEN *et al.*, 2008). Todas as imagens foram editadas no *software* Photoshop CS6. As cores originais foram mantidas, enquanto a região SCR foi destacada.

Posteriormente, a posição de ligação dos fatores de transcrição dentro da região SCR foi investigada. Para isso, o *software* Jaspar foi utilizado para gerar uma lista de pequenas sequências de nucleotídeos que possam ser a base para a ligação e interação de fatores de transcrição (MATHELIER *et al.*, 2013). Esses trechos foram selecionados comparando a sequência da região SCR com os motivos de ligação de DNA conhecidos para cada fator de transcrição. Com isso, foi possível plotar as possíveis posições de ligação de fatores de transcrição dentro da região SCR no mapa genômico da Universidade da Califórnia Santa Cruz em sua versão mm9 (KENT *et al.*, 2002). Complementarmente, os dados de ligação de fatores de transcrição foram obtidos através do consórcio ENCODE (MOUSE ENCODE CONSORTIUM *et al.*, 2012). Os sítios de ligação que se sobrepunham às áreas de alta intensidade de ligação de fatores de transcrição foram destacados. As imagens foram editadas no *software* Photoshop CS6.

4 RESULTADOS

4.1 PRIMEIRO BLOCO EXPERIMENTAL

4.1.1 Detecção da proteína Sox2 em diferentes clones celulares

Os efeitos que as deleções genômicas nas regiões reguladoras SRR1, SRR2, SRR18 e SCR exercem sobre a expressão da proteína Sox2 foram mensurados em células-tronco embrionárias.

A Figura 5 exibe os níveis da proteína Sox2 extraídas das amostras F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 30, 54 e 91), $\Delta\text{SRR1}/2/18^{\text{Cast}}$ (Clone 39) e $\Delta\text{SRR2-SRR104}^{129/\text{Cast}}$ (Clone 18 e 21). Um total de 7.9 μg de extrato proteico foi inserido em cada poço do gel. Enquanto os níveis proteicos de Gapdh estão aproximadamente estáveis entre as amostras, é possível observar variações nos níveis de Sox2 em cada clone. Células com deleções nas regiões reguladoras SRR1 + SRR2 + SRR18, SRR2, e SRR2 a SRR104 possuem flutuações nos níveis da proteína Sox2 comparados com o controle F1. Já a deleção homozigótica da região SCR resultou na privação completa da expressão de Sox2.

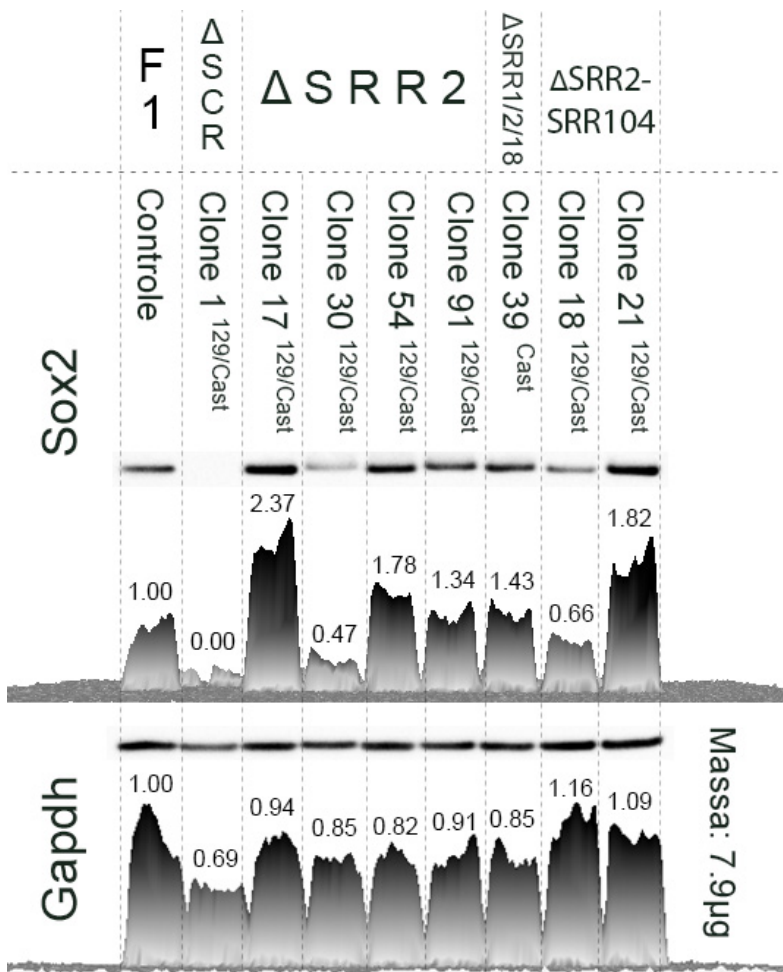
A Figura 6 mostra os níveis detectados da proteína Sox2, mas desta vez extraídas das amostras F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 1, Clone 4, Clone 30 e Clone 37), $\Delta\text{SRR18-104}^{\text{Cast}}$ (Clone 21 e 39) e $\Delta\text{SRR18-104}^{129/\text{Cast}}$ (Clone 70). Novamente, é possível observar que os níveis de Gapdh estão semelhantes entre os clones. Os clones celulares contendo as deleções nas regiões reguladoras SRR18 e SRR18 a SRR104 exibiram iguais ou menores níveis da proteína Sox2 comparados ao controle F1. Mais uma vez, o clone com deleção homozigótica da região SCR foi o único a exibir completa ausência da proteína Sox2.

A Figura 7 revela os níveis da proteína Sox2 extraídas das amostras F1 (Controle), F1 (Diferenciado), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18-104}^{\text{Cast}}$ (Clone 10), $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 62), $\Delta\text{SOX2}^{\text{Cast}}$ (Clone 36), $\Delta\text{SRR1}^{129/\text{Cast}} + \Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 25 e 39) e $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 36). É possível observar que a expressão de Sox2 é extremamente baixa em células-tronco embrionárias diferenciadas. Mais uma vez, os níveis de Sox2 variaram entre os diferentes clones com deleções nas regiões SRR18 a SRR104, SRR2, SRR1 + SRR18 e SRR18. Os níveis de Gapdh permaneceram invariáveis de forma geral. Porém, os clones provenientes do grupo de deleção $\Delta\text{SRR1}^{129/\text{Cast}} + \Delta\text{SRR18}^{129/\text{Cast}}$ possuíam problemas de crescimento, e por isso não foi possível obter a concentração proteica necessária. Nesse caso, todo o conteúdo obtido foi utilizado. Um cálculo da relação entre o volume proteico inserido e o volume padrão das outras amostras foi realizado, permitindo a melhor comparação entre amostras.

Somente o clone com deleção da região SCR e o clone F1 diferenciado exibiram completa falta da proteína Sox2.

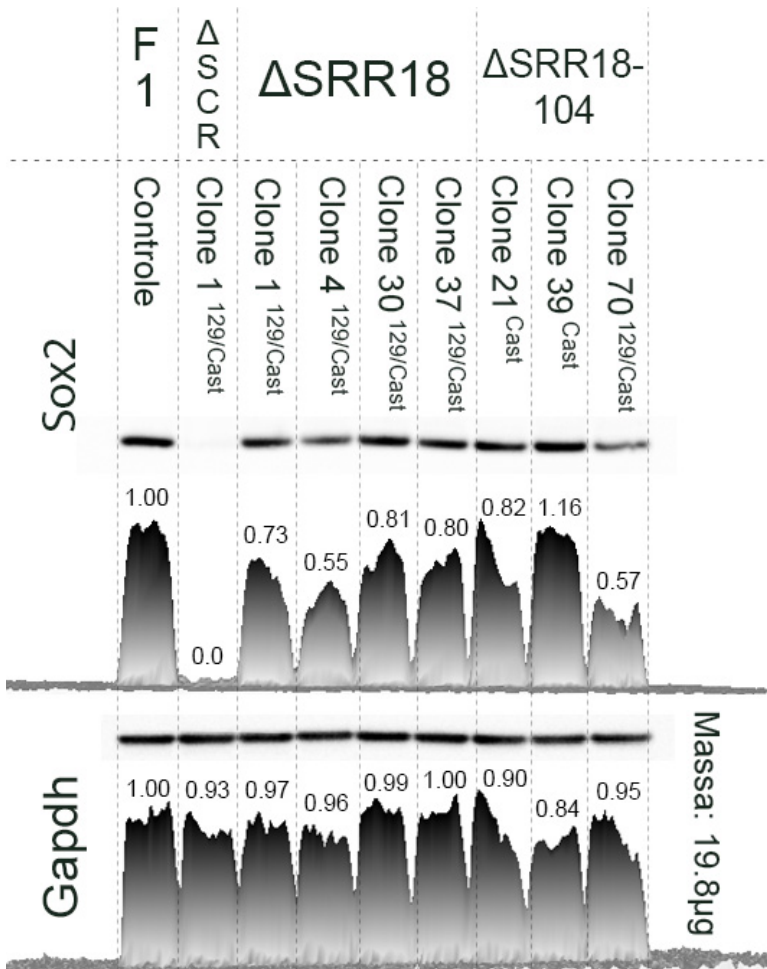
A Figura 8 demonstra a última imunodeteção realizada. Nesse caso, os clones que exibiram alguma variação inesperada nos géis anteriores foram colocados juntos e novamente analisados. Então, as amostras F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR18}^{129/\text{Cast}}$ (Clone 30) e $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 62, 54 e 17) tiveram seus níveis proteicos mensurados. Outra vez, os níveis de Gapdh permaneceram consistentes enquanto os níveis de Sox2 oscilaram dentro desses clones, quando comparados com a F1. Apesar das células com deleções na região SRR2 apresentarem uma diminuição da proteína Sox2, o único clone a exibir a ausência completa da proteína foi, novamente, o que possui deleção da região SCR.

Figura 5: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), $\Delta\text{SCR}^{129/\text{Cast}}$ (Clone 1), $\Delta\text{SRR2}^{129/\text{Cast}}$ (Clone 30, 54 e 91), $\Delta\text{SRR1/2/18}^{\text{Cast}}$ (Clone 39) e $\Delta\text{SRR2-SRR104}^{129/\text{Cast}}$ (Clone 18 e 21). A massa proteica total inserida em cada poço do gel foi de 7.9µg. Cada coluna representa um clone celular distinto, dentro de uma categoria de deleções. A concentração proteica é representada em sua forma original, em bandas, e em três dimensões (computadorizada). O valor número corresponde ao volume proteico aproximado em relação ao controle F1 e à proteína Gapdh. O símbolo $\Delta^{129/\text{Cast}}$ representa uma deleção em ambos os alelos, enquanto Δ^{Cast} representa uma deleção no alelo *Castaneus*.



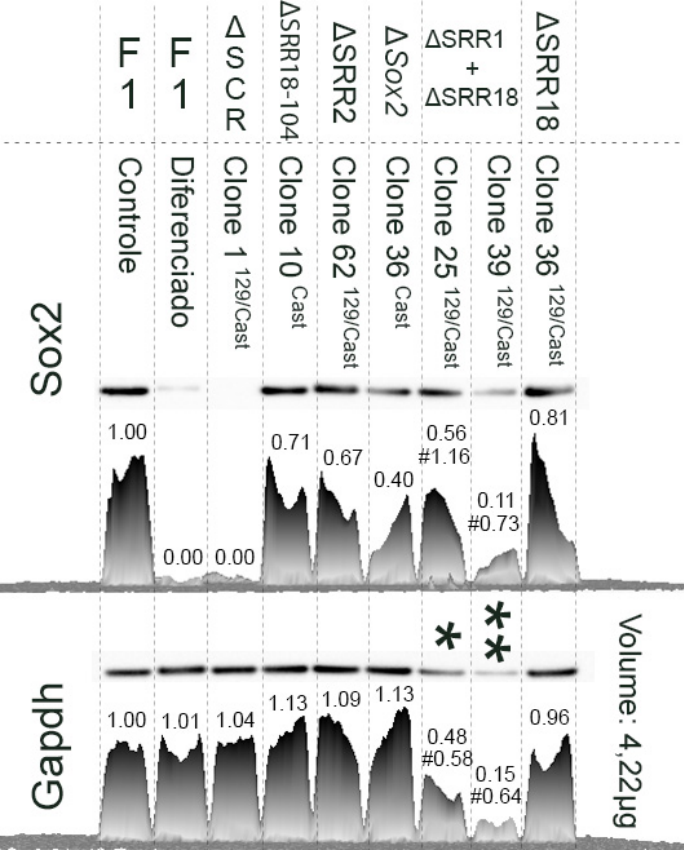
Fonte: O autor (2015).

Figura 6: Imunodeteção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), Δ SCR^{129/Cast} (Clone 1), Δ SRR18^{129/Cast} (Clone 1, Clone 4, Clone 30 e Clone 37), Δ SRR18-104^{Cast} (Clone 21 e 39) e Δ SRR18-104^{129/Cast} (Clone 70). A massa proteica total inserida em cada poço do gel foi de 19.8µg. Cada coluna representa um clone celular distinto, dentro de uma categoria de deleções. A concentração proteica é representada em sua forma original, em bandas, e em três dimensões (computadorizada). O valor número corresponde ao volume proteico aproximado em relação ao controle F1 e à proteína Gapdh. O símbolo Δ ^{129/Cast} representa uma deleção em ambos os alelos, enquanto Δ ^{Cast} representa uma deleção no alelo *Castaneus*.



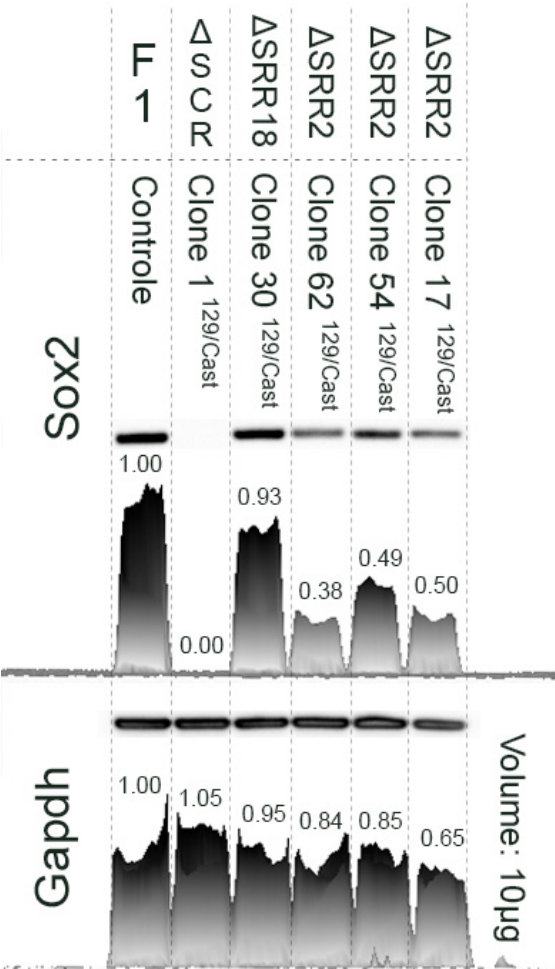
Fonte: O autor (2015).

Figura 7: Imunodeteccção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), F1 (Diferenciado), Δ SCR^{129/Cast} (Clone 1), Δ SRR18-104^{Cast} (Clone 10), Δ SRR2^{129/Cast} (Clone 62), Δ SOX2^{Cast} (Clone 36), Δ SRR1^{129/Cast} + Δ SRR18^{129/Cast} (Clone 25 e 39) e Δ SRR18^{129/Cast} (Clone 36). A massa proteica total inserida em cada poço do gel foi de 4.22 μ g. Clones provenientes do grupo de deleção Δ SRR1^{129/Cast}+ Δ SRR18^{129/Cast} receberam uma massa proteica menor (* = 3.52 μ g e ** = 1 μ g). Cada coluna representa um clone celular distinto, dentro de uma categoria de deleções. A concentração proteica é representada em sua forma original, em bandas, e em três dimensões (computadorizada). O valor número corresponde ao volume proteico aproximado em relação ao controle F1 e à proteína Gapdh. O símbolo # representa o volume aproximado relativizado à menor massa proteica inserida nas devidas amostras. O símbolo Δ ^{129/Cast} representa uma deleção em ambos os alelos, enquanto Δ ^{Cast} representa uma deleção no alelo *Castaneus*.



Fonte: O autor (2015).

Figura 8: Imunodeteccção da proteína Sox2 e da proteína controle Gapdh provenientes de células F1 (Controle), Δ SCR^{129/Cast} (Clone 1), Δ SRR18^{129/Cast} (Clone 30) e Δ SRR2^{129/Cast} (Clone 62, 54 e 17). A massa proteica total inserida em cada poço do gel foi de 10µg. Cada coluna representa um clone celular distinto, dentro de uma categoria de deleções. A concentração proteica é representada em sua forma original, em bandas, e em três dimensões (computadorizada). O valor número corresponde ao volume proteico aproximado em relação ao controle F1 e à proteína Gapdh. O símbolo Δ ^{129/Cast} representa uma deleção em ambos os alelos, enquanto Δ ^{Cast} representa uma deleção no alelo *Castaneus*.



Fonte: O autor (2015).

4.2 SEGUNDO BLOCO EXPERIMENTAL

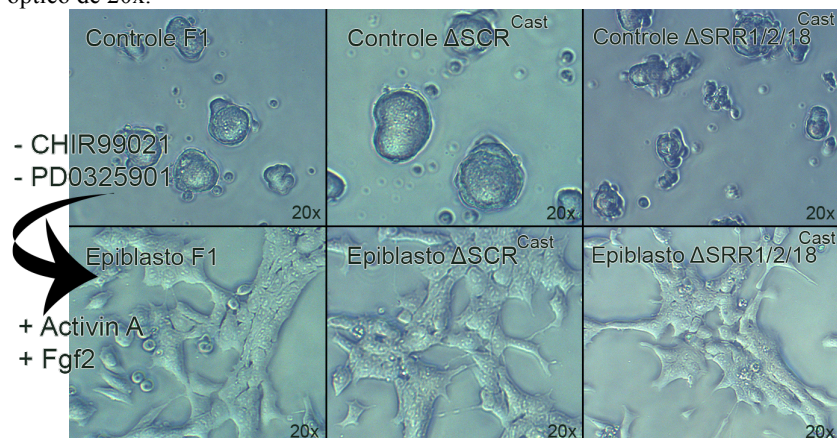
4.2.1 A diferenciação de células-tronco embrionárias

Para verificar se a região SCR continua exercendo um controle majoritário sobre a expressão de *Sox2* após a diferenciação celular, células-tronco embrionárias com deleções gênicas nas regiões reguladoras de *Sox2* foram induzidas ao estado de epiblasto. Para isso, células-tronco embrionárias de camundongo cultivadas em meio de cultura N2B27 controle contendo os inibidores de diferenciação CHIR99021 e PD0325901 foram transferidas para placas com meio de cultura N2B27 indutor de epiblasto incluindo Activin A e Fgf2. Todos os três grupos celulares, com deleções em regiões reguladoras específicas e células-tronco controle – F1, $\Delta\text{SCR}^{\text{Cast}}$, $\Delta\text{SRR1/2/18}^{\text{Cast}}$ – foram cultivados sob esse meio de cultura indutor por 48 horas.

Todas as células sofreram evidentes modificações morfológicas após esse tempo de cultivo sob o meio de cultura indutor de diferenciação, quando comparados com seus mesmos grupos que permaneceram no meio de cultura controle. Essa transformação pode ser observada na Figura 9. Células-tronco sob o meio de cultura N2B27 indutor exibiram sinais de diferenciação, como bordas irregulares e uma conformação achatada na placa de cultivo, enquanto células que permaneceram no meio de cultura N2B27 controle continuaram com um padrão celular arredondado e com formas suaves.

De modo geral, células-tronco dos três grupos experimentais resistiram bem à mudança de meio de cultura, não exibindo um aumento significativo de morte celular sob o meio N2B27 indutor. Entretanto, com o passar das gerações, células-tronco que permaneceram no meio de cultura N2B27 controle começaram a exibir um considerável aumento nos níveis de morte celular, com um grande número de células se despreendendo do fundo das placas. Ainda, o grupo de células $\Delta\text{SRR1/2/18}^{\text{Cast}}$ em tratamento controle exibiu problemas na taxa de divisão celular nas três réplicas biológicas, enquanto o grupo de células $\Delta\text{SCR}^{\text{Cast}}$ sob tratamento controle sofreu o mesmo fenômeno apenas na réplica biológica de número 3. Em ambos os casos, o menor número de células em cultura resultou em uma menor densidade celular final e, portanto, uma queda na qualidade e na concentração de RNA extraído. Esse fato pode ter influenciado nos resultados obtidos posteriormente.

Figura 9: Imagem de microscópio de luz invertida dos grupos celulares F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$ e de seus respectivos tratamentos, controle e induzido a epiblasto, após 48 horas de cultivo em seus meios de cultura N2B27 e N2B27 indutor de epiblasto. Todas as imagens foram capturadas utilizando aumento óptico de 20x.



Fonte: O autor (2015).

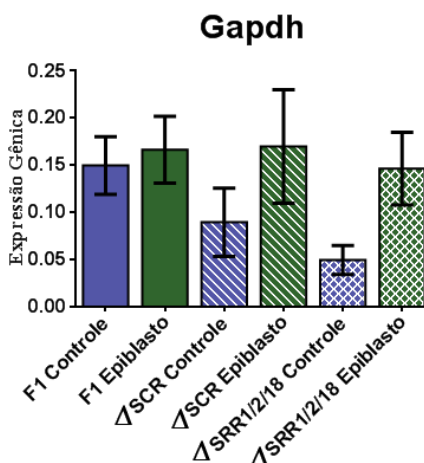
4.2.2 O padrão de expressão gênica de células-tronco induzidas a epiblasto

O perfil de expressão gênica de células-tronco induzidas a epiblasto foi obtido através de PCR quantitativo de transcrição reversa. Além do gene *Sox2*, os níveis de transcrição dos genes indicadores de estado celular *Dnmt3a*, *Esrrb*, *Klf4*, *Nanog*, *Oct4*, *Otx2*, *Pou3f1*, *Tbx3* e *Wnt8a* foram analisados nos três grupos celulares experimentais – F1 Controle, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$. Todos os valores de expressão gênica foram normalizados à expressão do gene controle *Gapdh*. O grupo de células da réplica de número 3 apresentou problemas na amplificação dos genes *Pou3f1* e *Wnt8a*, e por isso somente duas réplicas foram consideradas. Em todos os outros casos, três réplicas biológicas foram utilizadas nos cálculos de expressão gênica de cada grupo e tratamento.

4.2.2.1 A expressão de *Gapdh*

A expressão do gene *Gapdh* foi similar em todos os grupos celulares, exceto no tratamento controle das células $\Delta\text{SRR1/2/18}^{\text{Cast}}$, como pode ser visto na Figura 10. A baixa expressão de *Gapdh* em células $\Delta\text{SRR1/2/18}^{\text{Cast}}$ controle representa um baixo número de células em cultura, como mencionado anteriormente, e pode influenciar os resultados obtidos em outros experimentos.

Figura 10: Expressão do gene *Gapdh* obtida através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$. Os valores de expressão gênica representam a concentração de DNA complementar presente na amostra em relação à uma concentração de DNA conhecida. A barra representa o erro padrão amostral.

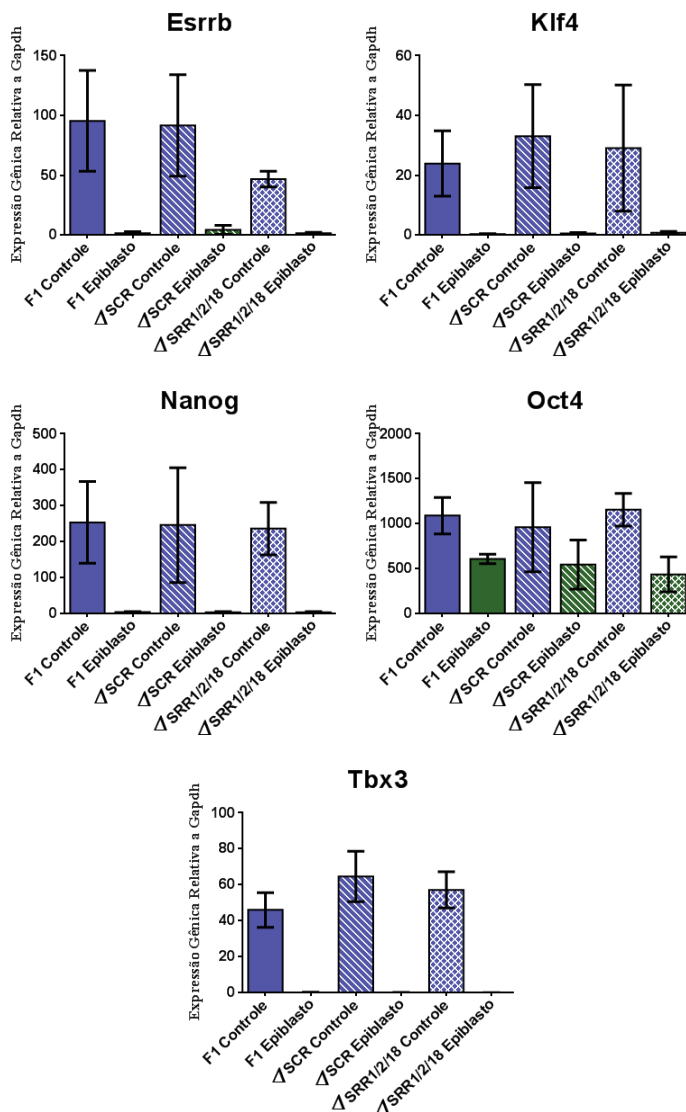


Fonte: O autor (2015).

4.2.2.2 Os genes *Esrrb*, *Klf4*, *Nanog*, *Oct4* e *Tbx3* possuem sua expressão reduzida em células-tronco induzidas a epiblasto

Cinco genes apresentaram flutuação negativa em seus níveis de expressão comparando células-tronco embrionárias controle e células-tronco induzidas a epiblasto, como representado na Figura 11. *Esrrb*, *Klf4*, *Nanog* e *Tbx3* exibiram uma grande diminuição de sua expressão perante a diferenciação das células-tronco embrionárias, enquanto o gene *Oct4* apresentou apenas uma sutil variação entre estados celulares. Curiosamente, a expressão de *Esrrb* proveniente de células $\Delta\text{SRR1/2/18}^{\text{Cast}}$ controle foi notavelmente menor que células F1 controle. Em nenhum desses resultados foi encontrado diferenças estatísticas. Porém, a diferença entre os valores de expressão de cada gene é notável entre tratamentos e grupos celulares.

Figura 11: Expressão dos genes *Esrrb*, *Klf4*, *Nanog*, *Oct4* e *Tbx3* obtidas através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$. Os valores de expressão gênica representam a concentração de DNA complementar presente na amostra normalizadas ao gene *Gapdh*. A barra representa o erro padrão amostral.



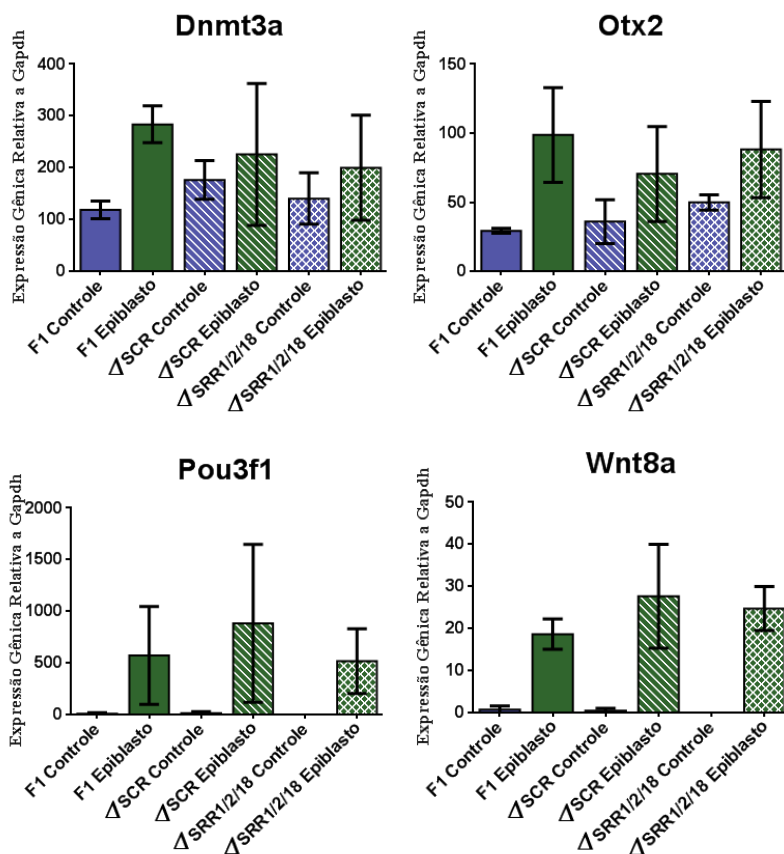
Fonte: O autor (2015).

4.2.2.3 Os genes *Dnmt3a*, *Otx2*, *Pou3fl* e *Wnt8a* apresentam uma expressão aumentada em células-tronco induzidas a epiblasto

Em contraste com os resultados anteriores, quatro genes apresentaram uma variação positiva em sua expressão após a diferenciação das células-tronco embrionárias em células-tronco induzidas a epiblasto, como pode ser visto na Figura 12. Enquanto *Pou3fl* e *Wnt8a* manifestaram um claro aumento de expressão perante a diferenciação celular, *Dnmt3a* e *Otx2* exibiram leves oscilações positivas em seus níveis de transcrição. Não houve diferenças estatísticas detectadas nesses resultados. Porém, parece claro que *Pou3fl* e *Wnt8a* possuem uma importante diferença de expressão entre os dois tratamentos já que os erros padrões respectivos não se sobrepõem. Já as diferenças na expressão de *Dnmt3a* e *Otx2* após a diferenciação celular parecem mais sutis.

É importante notar, porém, que a terceira réplica biológica de experimentos falhou em apresentar a expressão do gene *Pou3fl* e do gene *Wnt8a*, e por isso foi desconsiderada. Por esse motivo, os níveis de expressão de *Pou3fl* e *Wnt8a* apresentam um erro padrão maior que outros genes, já que somente representam duas réplicas biológicas.

Figura 12: Expressão dos genes *Dnmt3a*, *Otx2*, *Pou3f1* e *Wnt8a* obtidas através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, $\Delta\text{SCR}^{\text{Cast}}$ e $\Delta\text{SRR1/2/18}^{\text{Cast}}$. Os valores de expressão gênica representam a concentração de DNA complementar presente na amostra normalizadas ao gene *Gapdh*. A barra representa o erro padrão amostral. *Pou3f1* e *Wnt8a* apresentam apenas duas réplicas biológicas (n=2).



Fonte: O autor (2015).

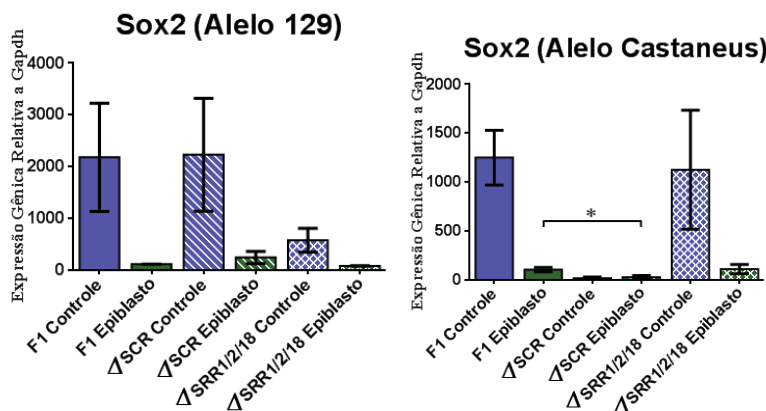
4.2.2.4 A influência da região SCR na expressão de *Sox2*

Por último, a expressão de *Sox2* proveniente dos alelos *129* – intacto – e alelo *Castaneus* – contendo as deleções – foi mensurada. Em ambos os alelos, foi encontrado um padrão de diminuição da expressão de *Sox2* perante a diferenciação de células-tronco embrionárias em células-tronco induzidas a epiblasto, como pode ser visto na Figura 13.

A expressão de *Sox2* do alelo *129* é alta no tratamento controle em todos os grupos celulares e diminui nas células-tronco sob tratamento de indução ao epiblasto. A exceção parece ficar com células-tronco embrionárias no meio controle com a deleção $\Delta\text{SRR1/2/18}^{\text{Cast}}$. Essas células demonstraram uma menor expressão de *Sox2* vinda do Alelo *129* quando comparadas com as células F1 Controle.

Já a expressão de *Sox2* do alelo *Castaneus* é diferente. Enquanto células-tronco F1 e $\Delta\text{SRR1/2/18}^{\text{Cast}}$ apresentaram grande expressão de *Sox2* sob o tratamento controle, essa expressão diminui de maneira semelhante sob o meio de indução ao epiblasto. Porém, células-tronco com a deleção $\Delta\text{SCR}^{\text{Cast}}$ apresentaram um grande impacto na impressão de *Sox2* tanto no tratamento controle, quanto no tratamento de indução ao epiblasto. O teste estatístico ANOVA considerou as diferenças de expressão gênica de *Sox2* do alelo *Castaneus* provenientes das células F1 e $\Delta\text{SCR}^{\text{Cast}}$ estatisticamente diferentes sob o tratamento de indução ao epiblasto.

Figura 13: Expressão do gene *Sox2* proveniente do alelo *129* e do alelo *Castaneus* obtida através de PCR quantitativo de três réplicas biológicas (n=3) nos tratamentos controle e induzidos a epiblasto de células F1, Δ SCR^{Cast} e Δ SRR1/2/18^{Cast}. Os valores de expressão gênica representam a concentração de DNA complementar presente na amostra normalizadas ao gene *Gapdh*. A barra representa o erro padrão amostral. Valores de $p < 0.05$ foram assinalados com * conforme teste ANOVA de uma via.



Fonte: O autor (2015).

4.3 TERCEIRO BLOCO EXPERIMENTAL

4.3.1 O papel molecular da região SCR

Experimentos *in silico* foram realizados de modo a investigar o papel molecular das regiões reguladoras de *Sox2* em células-tronco embrionárias. Sequenciamentos de imunoprecipitação de cromatina e de atividade da enzima DNase I obtidos *in silico* através do consórcio ENCODE permitem a verificação dos locais em que fatores de transcrição se ligam no genoma do camundongo, e em que nível de compactação esse genoma se apresenta em células-tronco embrionárias. O alinhamento dos resultados desses experimentos com as regiões reguladoras SRR1, SRR2, SRR18, SCR e com o gene *Sox2* através do mapa genômico da Universidade da Califórnia em Santa Cruz pode ser verificado na Figura 14. A área correspondente à região SCR foi destacada em cada tipo de experimento.

Primeiramente, é possível verificar que os picos que indicam as áreas de maior ligação de fatores de transcrição correspondem ao próprio gene *Sox2*, à região SRR2, e principalmente à região SCR. Nessa última

região, o tamanho dos picos é predominantemente maior que o de outras áreas do genoma próximas ao gene *Sox2*.

Similarmente, modificações de histonas se mostram mais significantes na região do gene *Sox2* e especialmente na região SCR, onde essas alterações químicas nas caudas de histonas parecem predominar.

Metilações na lisina 4 da cauda de histonas H3 (H3K4me1) parecem se concentrar nas regiões reguladoras SRR1, SRR2, SRR18 e principalmente na região SCR. É interessante notar, porém, que há uma região intermediária entre o gene *Sox2* e a região SCR em que não houve detecção de modificações do tipo H3K4me1.

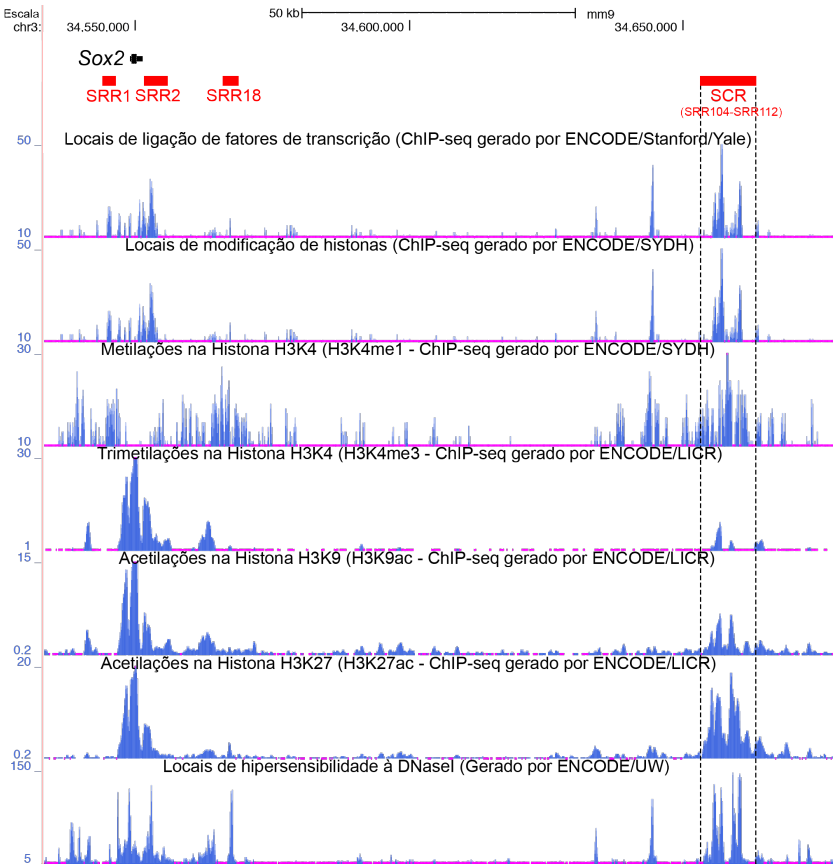
Já os picos de detecção de trimetilações na lisina 4 da cauda de histonas H3 (H3K4me3) se mostram prevalentes no gene *Sox2*, nas regiões SRR1, SRR2 e SRR18, continuamente. Porém, somente a região SCR apresenta um pico considerável da histona H3K4me3 em uma região mais afastada de qualquer gene, incluindo *Sox2*.

Esse padrão novamente se repete na detecção de acetilações na lisina 9 da cauda de histonas H3 (H3K9ac) e na lisina 27 da cauda de histonas H3 (H3K27ac). Mais uma vez, a região SCR se reforça como sendo a única região remota e distante do gene *Sox2* a exibir uma grande quantidade de modificações epigenéticas da histona H3.

Por fim, dados da atividade da enzima DNase I foram plotados no mapa genômico da UCSC *in silico*. Quanto maior a atividade da enzima DNase registrada, maior é o nível de acessibilidade da região genômica à fatores externos, isto é, menor é a sua compactação. É possível observar que embora as regiões próximas ao gene *Sox2* resultam em grande atividade dessa enzima, somente a região SCR exibe uma intensa atividade de ação da DNase I, mesmo não estando próximo de qualquer outro gene.

Uma região não caracterizada e cerca de 10.000 pares de base mais próxima do gene *Sox2*, em relação à região SCR, também apresentou picos de ligação de fatores de transcrição, de modificações H3Kme1 e de hipersensibilidade à DNase I.

Figura 14: Mapa genômico contendo informações das regiões reguladoras SRR1, SRR2, SRR18 e SCR. Experimentos de sequenciamento de imunoprecipitação de cromatina (ChIP-Seq) foram realizados pelo ENCODE e parceiros (MOUSE ENCODE CONSORTIUM *et al.*, 2012). Experimentos de DNase foram realizados pelo ENCODE e parceiros (MOUSE ENCODE CONSORTIUM *et al.*, 2012). Os picos representam áreas de maior atividade detectada. A escala de 50.000 pares de bases (kb) é representada abaixo. Essa ilustração foi gerada *in silico* através do mapa genômico da UCSC, utilizando-se a compilação do genoma de camundongo mm9 (KENT *et al.*, 2002).



Fonte: KENT *et al.* (2002)

4.3.2 Os fatores de transcrição relacionados à região SCR

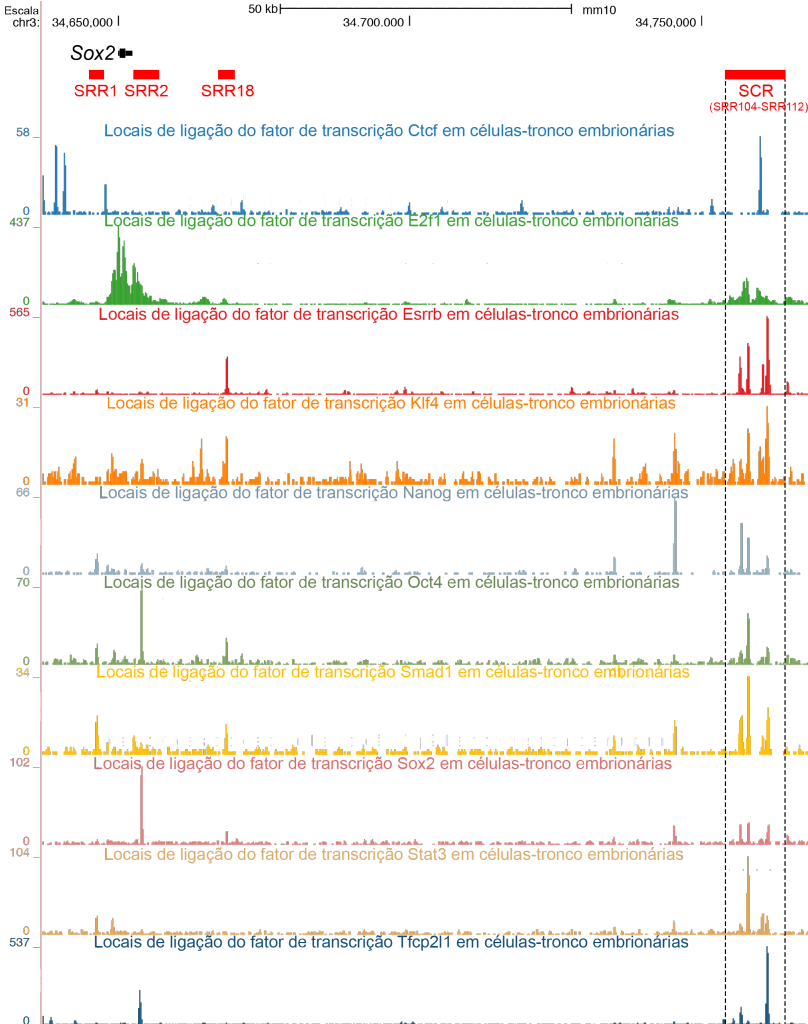
Para investigar quais fatores de transcrição possuem atividade de ligação nas regiões reguladoras de *Sox2*, o banco de dados CODEX foi consultado. A partir disso, experimentos de sequenciamento de imunoprecipitação de cromatina com diversos fatores de transcrição foram plotados no mapa genômico da UCSC *in silico*. Dessa maneira, o pico de atividade de fatores de transcrição e suas posições em relação as regiões reguladoras SRR1, SRR2, SRR18, SCR e ao gene *Sox2* podem ser vistos na Figura 15.

De maneira geral, todos os experimentos de imunoprecipitação encontrados com fatores de transcrição em células-tronco embrionárias exibiram um pico de ligação dentro da região SCR. Os fatores de transcrição Ctf, Esrrb, Klf4, Nanog, Smad1, Stat3 e Tfcp2l1 exibiram seus maiores picos de atividade na região SCR, enquanto Oct4 e Sox2 manifestaram picos maiores dentro da região SRR2, embora também apresentaram atividade na região SCR. E2f1 apresentou uma maior atividade na região do gene *Sox2*, apesar de também exibir uma leve ocupação na região SCR. Curiosamente, Klf4, Nanog, Smad1 e Sox2 manifestaram um pico de atividade em uma área desconhecida, fora da região SCR.

Posteriormente, o *software* Jaspar foi utilizado para apurar em que posições exatas fatores de ligação podem estar interagindo dentro da região SCR. Para isso, a sequência da região SCR foi analisada pelo *software*, e uma lista de trechos de nucleotídeos em que possíveis fatores de transcrição estejam se ligando foi retornada. Essa lista foi, então, plotada no mapa genômico da UCSC e comparada aos picos de ligação de fatores de transcrição obtidos através de experimentos de imunoprecipitação. Os resultados podem ser conferidos na Figura 16. A área de congruência entre os resultados de imunoprecipitação com a região SCR foi destacada.

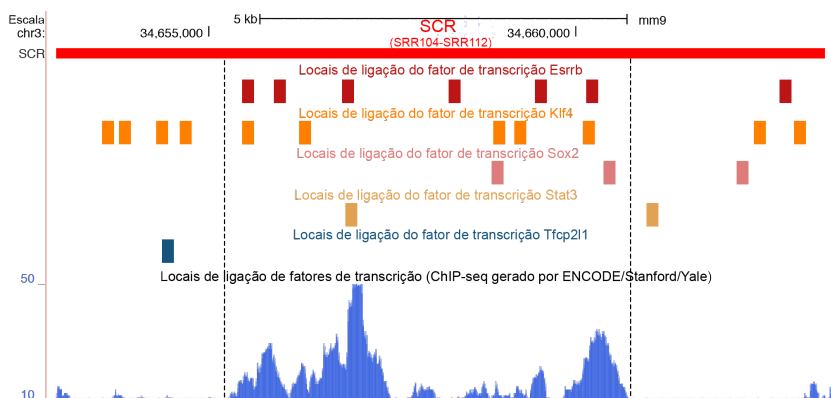
Jaspar encontrou sete possíveis sítios de ligação para o fator de transcrição Esrrb, seis deles dentro da região de grande atividade de fatores de transcrição indicadas pelo experimento de imunoprecipitação. Klf4, por sua vez, resultou em onze possíveis sítios de ligação, sendo cinco deles dentro da região congruente. Sox2 retornou três sítios de ligação, com dois deles dentro da região assinalada. Dois sítios de ligação de Stat3 foram indicados, sendo um deles dentro da região destacada. E, por fim, Tfcp2l1 retornou somente um possível sítio de ligação, deslocado da área de maior atividade de fatores de transcrição na região SCR.

Figura 15: Mapa genômico contendo dez fatores de transcrição com atividade de ligação nas áreas reguladoras SRR1, SRR2, SRR18 e SCR. Os picos representam áreas de maior interação detectada. Esses dados foram obtidos através de experimentos de sequenciamento de imunoprecipitação de cromatina (ChIP-Seq) realizados por Chen *et al.* (2008). A escala de 50.000 pares de bases (kb) é representada abaixo. Essa ilustração foi gerada *in silico* através do mapa genômico da UCSC, utilizando-se a compilação do genoma de camundongo mm10 (KENT *et al.*, 2002).



Fonte: KENT *et al.* (2002)

Figura 16: Possíveis locais de ligação dos fatores de transcrição *Esrrb*, *Klf4*, *Sox2*, *Stat3* e *Tfcp2l1* na região SCR conforme resultados obtidos *in silico* pelo software Jaspar (MATHÉLIER *et al.*, 2013). Os resultados foram comparados com dados de sítios de ligação de fatores de transcrição, obtidos através do ENCODE e parceiros (MOUSE ENCODE CONSORTIUM *et al.*, 2012). A área dentro da região SCR que contém picos de ligação de fatores de transcrição foi destacada. A escala de 50.000 pares de bases (kb) é representada abaixo. Essa ilustração foi gerada através do mapa genômico da UCSC, utilizando-se a compilação do genoma de camundongo mm9 (KENT *et al.*, 2002).



Fonte: O autor (2015).

5 DISCUSSÃO

O controle dos fatores de transcrição que governam a manutenção e diferenciação de células-tronco embrionárias é de grande interesse dentro do campo científico. Isso ocorre principalmente devido às possibilidades de criação e utilização clínica de terapias de cultivo e regeneração celular, além da compreensão do processo de origem e desenvolvimento tumoral, por exemplo (NIH, 2015). No presente trabalho, foram estudados os mecanismos de controle da expressão do fator de transcrição *Sox2* em células-tronco embrionárias, bem como após o processo de diferenciação em células-tronco induzidas a epiblasto. Ao total, quatro regiões reguladoras apresentaram certa influência sobre a expressão do gene *Sox2*: a região SRR1, SRR2, SRR18, e principalmente a região SCR.

A região SCR foi caracterizada como sendo a principal região reguladora responsável pela expressão de *Sox2* em células-tronco embrionárias. Isso foi comprovado através da imunodeteção da proteína *Sox2* em clones celulares com diferentes deleções gênicas no primeiro bloco experimental. Somente a ausência da região SCR em ambos os alelos resultou na escassez completa de *Sox2*, enquanto deleções em outras regiões não foram capazes de apresentarem o mesmo efeito. De forma similar, a remoção da região SCR em ambos os alelos também já foi descrita por impactar a expressão de *Sox2* em células-tronco embrionárias (LI *et al.*, 2014; ZHOU *et al.*, 2014). Por outro lado, também foi detectada uma diminuição na proteína *Sox2* resultante da deleção de outras regiões reguladoras, como SRR1, SRR2 e SRR18. Isso pode representar possíveis influências que essas regiões também exerçam sobre o gene *Sox2*, embora menores que a região SCR. Além disso, diferentes clones contendo deleções na mesma região reguladora mostraram variações nos níveis proteicos de *Sox2*, provavelmente devido à diferença no tamanho da deleção genômica realizada pelo sistema CRISPR/Cas9 (MARUYAMA *et al.*, 2015). No entanto, nenhum desses clones exibiu o mesmo resultado obtido pela remoção da região SCR.

Para estudar a influência da região SCR no controle da expressão de *Sox2* após a diferenciação de células-tronco embrionárias, um modelo de indução ao epiblasto estabelecido por Hayashi *et al.* (2011) e Buecker *et al.* (2014) através da remoção dos inibidores CHIR99021 e PD032590, seguido de tratamento com Activin A e Fgf2 foi utilizado durante o segundo bloco experimental (HAYASHI *et al.*, 2011; BUECKER *et al.*, 2014). Células-tronco derivadas do epiblasto são caracterizadas por apresentarem um tamanho celular maior e colônias achatadas e compactadas. Células-tronco embrionárias, por outro lado, apresentam

características morfológicas distintas, como células arredondadas e colônias agrupadas em formas globulares (BRONS *et al.*, 2007; TESAR *et al.*, 2007). Após o tratamento de diferenciação, as mesmas características visuais compatíveis com células-tronco do epiblasto foram encontradas. A remoção dos inibidores CHIR99021 e PD032590 já é conhecida por resultar na diferenciação de células-tronco embrionárias (YING *et al.*, 2008). Como Activin A e Fgf2 são fatores de crescimento utilizados para a manutenção de células-troncos do epiblasto *in vitro* (BRONS *et al.*, 2007), é possível que sua administração tenha resultado na diferenciação de células-tronco embrionárias em células-tronco induzidas a epiblasto.

Para confirmar molecularmente essa diferenciação, nove genes indicadores de estado celular tiveram seu padrão de expressão analisados, antes e depois da administração de Activin A e Fgf2. A expressão do gene *Gapdh* foi utilizada para a padronização dos níveis de expressão entre grupo celulares. Nesse caso, somente células-tronco embrionárias do grupo celular Δ SRR1/2/18^{Cast} apresentaram níveis transcricionais menores de *Gapdh*. Provavelmente isso foi resultado da baixa densidade celular obtida durante os experimentos. Assim, com um menor número de células vivas, uma menor quantidade de transcrito total foi detectada. A expressão de outros genes provinda dessas células pode ter sido afetada pelo mesmo problema.

Buecker *et al.* (2014) relataram a diminuição da expressão dos genes *Esrrb*, *Klf4*, *Nanog*, *Oct4* e *Tbx3* após a diferenciação de células-tronco embrionárias em células-tronco induzidas a epiblasto (BUECKER *et al.*, 2014). De forma semelhante, os resultados aqui obtidos também indicam a diminuição da expressão desses genes, embora não haja uma diferença estatística clara entre os níveis transcricionais de cada estado celular. Como esses fatores de transcrição atuam em conjunto (BOYER *et al.*, 2005), e a diminuição de sua expressão já foi correlacionada com a diferenciação celular (IVANOVA *et al.*, 2006), esse parece ser mais um indicativo da ocorrência da indução ao epiblasto. Por isso, a ausência da diferença estatística na expressão desses genes muito provavelmente se relaciona ao baixo número de réplicas biológicas utilizadas nesse estudo, já que a análise ANOVA leva em consideração o número amostral para a análise das variâncias. *Oct4* foi o único gene que ainda era altamente expresso em células-tronco induzidas a epiblasto. Níveis de expressão de *Oct4* equivalentes aos de células-tronco embrionárias também foram detectados em células-tronco induzidas a epiblasto de outros estudos (HAYASHI *et al.*, 2011; JOO *et al.*, 2014). É possível que o fator de

transcrição *Oct4* ainda seja necessário na manutenção de células-tronco induzidas a epiblasto, o que explicaria sua constante expressão.

Os genes *Dnmt3a*, *Otx2*, *Pou3f1* e *Wnt8a* tiveram um aumento em seus níveis transcricionais após a diferenciação em células-tronco induzidas a epiblasto. A análise dessas células por Buecker *et al.* (2014) também exibiu os mesmos padrões de aumento de expressão desses genes após o tratamento de indução ao epiblasto (BUECKER *et al.*, 2014). Por isso, embora o teste ANOVA mais uma vez não tenha acusado diferença estatística entre os níveis transcricionais em cada tratamento celular, esses resultados reforçam a hipótese de que as células-tronco embrionárias foram diferenciadas em células-tronco induzidas a epiblasto. É também possível que haja uma correlação direta entre a diminuição da expressão dos genes anteriormente citados com o aumento da transcrição desses fatores. Por exemplo, *Dnmt3a* é responsável pela metilação do DNA (OKANO *et al.*, 1999). Porém, a expressão de *Tbx3* parece afetar essa metilação através da diminuição da expressão de *Dnmt3a* em células-tronco embrionárias (DAN *et al.*, 2013). Desse modo, a diminuição da expressão de *Tbx3* pode ter levado ao aumento da expressão de *Dnmt3a*. Com o aumento de *Dnmt3a*, *Oct4* e *Nanog* podem ter tido sua expressão afetada, já que o complexo de metilação Dnmt3 foi relatado como sendo um dos responsáveis pela inativação dos promotores de *Oct4* e *Nanog* durante o desenvolvimento embrionário (LI *et al.*, 2007).

Otx2, no entanto, possui um papel mais definido em células-tronco induzidas a epiblasto. Sua expressão, influenciada por *Fgf2*, parece ser contrária à expressão de *Nanog* e *Klf4*, isto é, com a diminuição da expressão de *Nanog* e *Klf4*, maiores níveis do transcrito de *Otx2* são detectados (ACAMPORA; GIOVANNANTONIO; SIMEONE, 2012). Dessa maneira, *Fgf2* pode ter levado ao leve aumento de *Otx2* que, por sua vez, contribuiu para a diminuição dos níveis de expressão dos genes relacionados à pluripotência celular, como *Klf4* e *Nanog*. A alta dos níveis de transcrição de *Otx2* também foi previamente relacionada com o aumento da expressão de *Pou3f1*, principalmente na formação dos progenitores neurais (ZHU *et al.*, 2014). Embora células-tronco induzidas a epiblasto ainda não estejam engajadas na diferenciação da neuroectoderme, o aumento da expressão de *Pou3f1* pode indicar o preparo para a diferenciação neural de parte das células do epiblasto.

O gene *Pou3f1* também foi relatado na interferência da expressão de *Wnt8a* (IWAUCHI-DOI *et al.*, 2012). É possível que ambos os aumentos de *Pou3f1* e *Wnt8a* ocorreram porque parte das células estariam se preparando para a diferenciação em progenitoras neurais, e outra parte estivessem se preparando para a diferenciação para mesoderme e células

da linha primitiva, principal função do gene *Wnt8a* (LIU *et al.*, 1999; LINDSLEY *et al.*, 2006). A interação entre Fgf e *Wnt8a* na formação da linha primitiva também já foi relatada em outros organismos (HARDY *et al.*, 2011). Isso pode levar a uma correlação entre a administração de Fgf2 e o aumento da expressão de *Wnt8a* em células-tronco induzidas a epiblasto. Embora uma das réplicas biológicas também falhou em demonstrar expressão de *Wnt8a*, os resultados obtidos parecem demonstrar os níveis esperados de expressão de *Wnt8a* nas células. Com esses dados, é plausível sugerir que células-tronco embrionárias expostas à Activin A e Fgf2 apresentam um perfil de expressão gênica similar ao do epiblasto.

Com a confirmação do estado induzido a epiblasto, torna-se possível analisar mais de perto o perfil de expressão de *Sox2* nesse novo estado celular. Os níveis de transcrição do gene *Sox2* provenientes do alelo 129 apresentaram uma redução perante a diferenciação de células-tronco embrionárias. De fato, a remoção de *Sox2* resulta na diferenciação de células-tronco embrionárias (IVANOVA *et al.*, 2006), e a indução ao estado de epiblasto apresenta um menor nível de transcrição de *Sox2* (HAYASHI *et al.*, 2011; BUECKER *et al.*, 2014). Dessa maneira, o alelo 129 – que estava intacto em todos os grupos celulares – mostrou um perfil de expressão de *Sox2* condizente com a diminuição da pluripotencialidade em células-tronco induzidas a epiblasto, mesmo que a diferença entre tratamentos não seja estatisticamente significativa. Inesperadamente, o grupo celular $\Delta\text{SRR1/2/18}^{\text{Cast}}$ exibiu menores níveis de expressão de *Sox2* em células-tronco embrionárias, quando comparado ao controle. Isso pode ter sido resultado da baixa densidade celular obtida durante a condução dos experimentos.

A expressão de *Sox2* proveniente do alelo *Castaneus* apresentou um perfil de expressão mais interessante. Enquanto os grupos celulares controle F1 e $\Delta\text{SRR1/2/18}^{\text{Cast}}$ tinham grande expressão de *Sox2* em células-tronco embrionárias, e exibiram similar diminuição após a diferenciação celular, o grupo $\Delta\text{SCR}^{\text{Cast}}$ – sem a região SCR – apresentou uma expressão extremamente baixa de *Sox2* desse alelo. A quantificação dessa expressão foi similar àquela encontrada antes do processo de diferenciação nesse grupo celular. Esses resultados permitem retirar algumas conclusões. Primeiramente, a região SCR parece continuar exercendo fundamental controle sob a expressão de *Sox2* em células-tronco induzidas a epiblasto. Embora outras regiões reguladoras também estivessem ausentes em seus respectivos grupos celulares, a região SCR demonstra ser a única região capaz de diminuir profundamente a

transcrição de *Sox2* em células-tronco embrionárias e células-tronco induzidas a epiblasto.

Nessas condições, é possível que a remoção da região SCR elimine a habilidade de recrutar fatores de transcrição que permitam a expressão de *Sox2*, resultando em sua depleção transcricional. Isso não ocorre com a retirada de outras regiões reguladoras. Porém, a remoção da região SRR1 já é conhecida por resultar na má-formação cerebral de camundongos (FERRI *et al.*, 2004), enquanto a região SRR2 é conhecida por interagir com o fator de transcrição Oct4 e com o próprio *Sox2* (TOMIOKA *et al.*, 2002). Seguindo essa linha de raciocínio, é possível que essas regiões reguladoras assumam uma maior função no controle de *Sox2* em níveis mais profundos de diferenciação celular, como durante a formação do sistema nervoso central.

Com a confirmação de que a região SCR é a região primordial no controle da expressão de *Sox2* em células-tronco embrionárias e células-tronco induzidas a epiblasto, é necessário entender como essa região funciona molecularmente. De forma geral, regiões reguladoras são um grande centro de ancoragem de fatores de transcrição (SPITZ; FURLONG, 2012) e possuem uma variedade de modificações de cauda de histona (CALO; WYSOCKA, 2013; SHLYUEVA; STAMPFEL; STARK, 2014). Através de análises *in silico* no terceiro bloco experimental, foi possível observar que a região SCR apresenta ambas características.

As metilações na lisina 4 da cauda de histonas H3 (H3K4me1) são reconhecidas por representarem áreas reguladoras (HEINTZMAN *et al.*, 2007). Já acetilações na lisina 27 da cauda de histonas H3 (H3K27ac) determinam a ativação dessas regiões (CREYGHTON *et al.*, 2010). Ambas essas modificações foram detectadas em altos níveis na região SCR. Trimetilações na lisina 4 da cauda de histonas H3 (H3K4me3), por outro lado, são geralmente associadas com regiões promotoras. Analogamente, a região mais próxima do gene *Sox2* exibiu grandes níveis dessa modificação. Porém, a região SCR também exibiu um certo nível dessa modificação em suas histonas. A alta proporção de H3K4me1 para H3K4me3 parece diferenciar uma região promotora de uma região reguladora (HEINTZMAN *et al.*, 2007). A presença de H3K4me3 também pode representar a atividade do complexo de transcrição da RNA Polimerase II e a consequente atividade dessa região reguladora sob o seu gene alvo (KOCH; ANDRAU, 2011).

Por fim, a presença de acetilações na lisina 9 da cauda de histonas H3 (H3K9ac) é correlacionada com regiões reguladoras ativas, juntamente com H3K27ac (KARMODIYA *et al.*, 2012). De tal forma, um

pico de ambas modificações foi encontrado na região SCR. Juntos, esses resultados parecem indicar que a região SCR possui, de fato, marcações moleculares que permitem sua caracterização como uma região reguladora.

Já a sensibilidade à ação da enzima DNase I é conhecida por indicar um maior nível de acessibilidade cromossômica (GROSS; GARRARD, 1988). Regiões reguladoras normalmente exibem uma maior sensibilidade a essa nuclease (LI, 2002). Assim, os picos de hipersensibilidade à DNase encontrados na região SCR parecem correlacioná-la à sua função de região reguladora. Esse maior nível de acessibilidade corrobora os resultados de modificações de histona obtidos anteriormente, confirmando que esta região está mesmo acessível à ligação de fatores de transcrição e que muito provavelmente está molecularmente ativa em células-tronco embrionárias.

É importante salientar que uma pequena região anterior à região SCR também exibiu picos de modificação de histona, hipersensibilidade à DNase e ligação de fatores de transcrição. Porém, a remoção dessa região, que se localiza entre a região SRR2 e SCR, não resultou na ausência da proteína Sox2 durante os experimentos de imunodeteção. Isso permite levantar a hipótese de que essa região possui uma outra função biológica diferente da região SCR. Seria interessante estudá-la mais a fundo para verificar se de fato ela não apresenta nenhum papel na expressão de Sox2.

Com base nas características reguladoras que a região SCR exhibe, os fatores de transcrição que se ligam a ela são interessantes alvos para estudo. Nesse cenário, novas análises *in silico* mostraram que os fatores de transcrição Ctf, E2f1, Esrrb, Klf4, Nanog, Oct4, Smad1, Stat3, Tcf21 e do próprio Sox2 interagem com a região SCR. Esse resultado concorda com o mecanismo previamente apresentado onde os fatores responsáveis pela pluripotencialidade celular atuam em conjunto, com cada um dos fatores autorregulando-se simultaneamente (BOYER *et al.*, 2005). O fator de transcrição Ctf, por outro lado, é conhecido por auxiliar na formação de estruturas cromossômicas que permitem o contato entre regiões reguladoras e seus genes alvos (PHILLIPS; CORCES, 2009). Dessa forma, é possível que o pico de atividade de Ctf encontrado dentro da região SCR indique a interação entre essa região e o gene Sox2. Já o fator de transcrição E2f1 é conhecido em sua associação com regiões promotoras (BIEDA *et al.*, 2006). Por esse motivo, o seu maior nível de atividade ao redor do gene Sox2 não surpreende. Devido à sua associação com o complexo de transcrição da RNA Polimerase II, é possível que sua

atividade na região SCR seja ligada ao acoplamento do complexo de transcrição quando em contato com o gene *Sox2*.

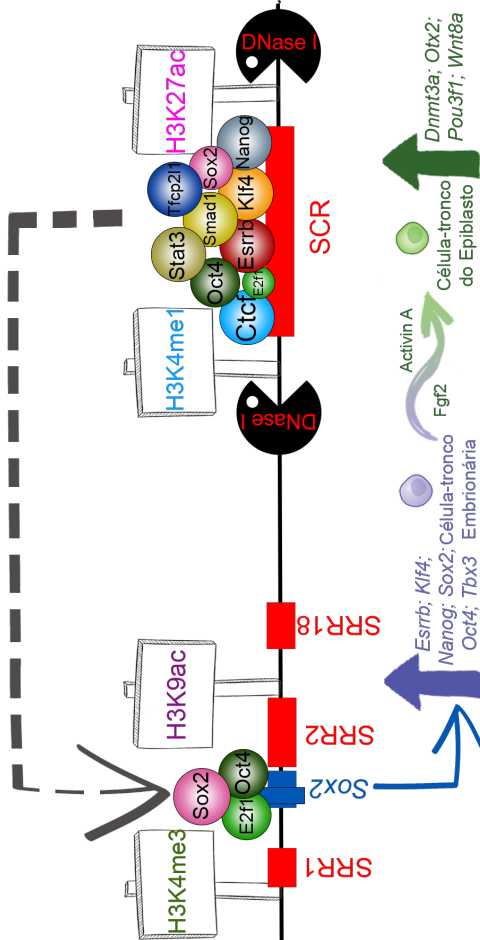
Os fatores de transcrição Smad1 e Stat3 parecem estar relacionados com Oct4 na manutenção da pluripotencialidade celular (CHEN *et al.*, 2008). De fato, os picos de atividade desses três fatores de transcrição estão localizados nas mesmas posições do genoma. Além disso, a ativação de Stat3 através da adição da molécula LIF (NIWA *et al.*, 1998) e de Smad1 através das proteínas morfogênicas ósseas (YING *et al.*, 2003) permitem a manutenção de células-tronco embrionárias *in vitro*. Ainda, a atividade de Tfcp2l1 foi relacionada com outros fatores de transcrição responsáveis pela pluripotencialidade celular, como Klf4, Nanog, Oct4 e Sox2 (YE *et al.*, 2013). Juntos, esses dados indicam que diversos fatores de transcrição atuam na região SCR de modo a permitir a manutenção da pluripotencialidade de células-tronco embrionárias através da expressão de *Sox2*. Já que a região SCR continua estimulando a expressão de *Sox2* em células-tronco induzidas a epiblasto, é provável que esta rede de controle, embora se modifique um pouco, ainda esteja presente e funcione de maneira análoga à apresentada nesse trabalho.

Se há tantos fatores de transcrição se ligando à região SCR, é muito provável que existam motivos de DNA dentro dessa região que permitam a ancoragem desses fatores de transcrição. Para analisar a existência desses possíveis motivos de DNA, o *software* Jaspar foi utilizado para verificar sua existência. A pouca quantidade de motivos de DNA encontrado em relação ao total de fatores de transcrição que se ligam na região SCR pode ser resultado do baixo número de motivos encontrados no banco de dados do Jaspar. De qualquer maneira, Esrrb, Klf4, Sox2 e Stat3 apresentaram motivos de ligação dentro da região SCR, corroborando os resultados previamente encontrados nesse estudo. Isso também parece confirmar que o fator de transcrição Sox2 possui um padrão de autocontrole transcricional (BOYER *et al.*, 2005; KIM *et al.*, 2008). O fator de transcrição Tfcp2l1 somente exibiu um sítio de ligação fora da região SCR, contrariando seu pico de atividade anteriormente descrito. Outros fatores que podem se ligar a essa região também não foram detectados por essa metodologia. Contudo, a ligação de Esrrb, Klf4, Sox2 e Stat3 dentro da região SCR coincide com os picos de ligação de fatores de transcrição detectados por sequenciamento de imunoprecipitação (CHEN *et al.*, 2008). Assim, embora superficiais, esses resultados concordam com a hipótese de que a região SCR é uma região reguladora que comporta a ligação de diferentes fatores de transcrição. Uma análise mais profunda desses motivos de DNA através

da mutação de suas sequências poderia levar a conclusões mais exatas de suas reais localizações.

O presente trabalho apresentou diferentes resultados que indicam que a expressão do gene *Sox2* parece ser controlada primordialmente pela região reguladora SCR em células-tronco embrionárias. Entretanto, outras regiões, como SRR1, SRR2, SRR18 podem ter diferentes influências no controle do gene em períodos tardios do desenvolvimento. Com a exposição de células-tronco embrionárias aos fatores de crescimento Activin A e Fgf2, essas células adquirem características condizentes com células-troncos extraídas do epiblasto. Por esse motivo, elas adotam um perfil celular conhecido como células-tronco induzidas a epiblasto. Esse novo estado celular foi caracterizado por exibir menores níveis de expressão de genes ligados à pluripotencialidade celular, como *Esrrb*, *Klf4*, *Nanog*, *Oct4* e *Tbx3*. Por outro lado, genes ligados ao desenvolvimento celular do epiblasto, como *Dnmt3a*, *Otx2*, *Pou3f1* e *Wnt8a* apresentaram um aumento em seus níveis transcricionais. Aqui, o gene *Sox2* também teve sua expressão diminuída. Porém, somente a deleção da região SCR no alelo *Castaneus* resultou em grande impacto na expressão de *Sox2* proveniente desse alelo em ambas células-tronco embrionárias e células-tronco induzidas a epiblasto. Esse efeito reforça a hipótese de que a região SCR é a principal região reguladora de *Sox2*, mesmo após a diferenciação celular. Mais a fundo, a região SCR foi caracterizada por apresentar diversas características cromossômicas similares a regiões reguladoras, como modificações de histona do tipo H3K4me3, H3K9ac, e principalmente H3K4me1 e H3K27ac. Além disso, um grande nível de hipersensibilidade a nucleases foi detectada, indicando um elevado nível de acessibilidade cromossômica. Essa acessibilidade se traduz na possibilidade de ligação de fatores de transcrição nessa região, como Ctf, E2f1, Esrrb, Klf4, Nanog, Oct4, Smad1, Sox2, Stat3 e Tfcp2l1. Essas proteínas, por sua vez, têm direta ou indireta ligação com a manutenção da pluripotencialidade celular. Juntos, esses resultados mostram que a região reguladora SCR é uma plataforma de ligação de fatores de transcrição que permite a expressão de *Sox2* e a manutenção da pluripotencialidade de células-tronco embrionárias. Esse efeito parece continuar sendo exercido em células-tronco induzidas a epiblasto. Um resumo dos mecanismos encontrados nesse trabalho pode ser visto na Figura 17.

Figura 17: Mecanismos de controle da expressão de *Sox2* em células-tronco embrionárias. As regiões reguladoras SRR1, SRR2, SRR18 e SCR estão representadas em vermelho. Os fatores de transcrição estão representados em formas esféricas em suas respectivas áreas de atuação. O tamanho das esferas corresponde ao seu pico de atividade em cada área. As modificações de histona presentes em cada região estão indicadas em suas respectivas placas. A região SCR possui hipersensibilidade à DNase I. Os fatores de transcrição se ligam na região SCR e levam à expressão de *Sox2*. *Sox2*, junto com outros fatores de transcrição, contribui para o estado pluripotente de células-tronco embrionárias. Células-tronco embrionárias são induzidas ao estado de epiblasto através de Activin A e Fgf2. Células-tronco induzidas a epiblasto exibem uma maior transcrição de diferentes fatores de transcrição.



Fonte: O autor (2015).

6 CONCLUSÕES

- A região SCR é a principal região reguladora responsável pela expressão de *Sox2* em células-tronco embrionárias, já que somente sua eliminação resultou na ausência da proteína Sox2.
- A utilização dos fatores de crescimento Activin A e Fgf2, e a remoção dos inibidores de diferenciação CHIR99021 e PD0325901, possibilitam a diferenciação de células-tronco embrionárias em células-tronco com perfil condizente ao de epiblasto.
- Células-tronco induzidas a epiblasto apresentam uma menor expressão dos genes *Esrrb*, *Klf4*, *Nanog*, *Oct4*, *Sox2* e *Tbx3*. Por outro lado, há um aumento da expressão de *Dnmt3a*, *Otx2*, *Pou3f1* e *Wnt8a*.
- A região SCR continua sendo a região mais importante para a expressão de *Sox2* em células-tronco induzidas a epiblasto, já que sua remoção no alelo *Castaneus* resulta em grande perda da expressão de *Sox2* proveniente desse alelo. As deleções feitas nas regiões SRR1, SRR2 e SRR18 não produzem o mesmo efeito.
- A região SCR possui características de região reguladora, como modificações de cauda de histonas e acessibilidade à ligação de fatores de transcrição. Esses fatores de transcrição são relacionados à manutenção da pluripotencialidade de células-tronco embrionárias e interagem com a região SCR através de alguns motivos de DNA.
- A região SCR é a principal região reguladora de *Sox2* em células-tronco embrionárias e células-tronco induzidas a epiblasto.

7 PERSPECTIVAS

No presente trabalho, foi possível caracterizar a região SCR como tendo profunda relação no controle da expressão de *Sox2* em células-tronco embrionárias e células-tronco induzidas a epiblasto. Porém, devido ao baixo número de réplicas biológicas, não foi possível mostrar grandes diferenças estatísticas entre o perfil de expressão de células-tronco embrionárias e de células-tronco induzidas a epiblasto. Dessa maneira, novos estudos contendo um maior número amostral poderiam aumentar a confiabilidade dos resultados aqui encontrados.

Ainda não está claro se o papel de controle da região SCR continua durante o desenvolvimento embrionário. Seria importante acompanhar o funcionamento do controle de *Sox2* durante as fases posteriores de diferenciação celular, bem como após a gastrulação. Isso se tornaria mais interessante na comparação de células que se diferenciem para os progenitores neurais e células que contribuam para outros folhetos embrionários, já que *Sox2* continua sendo expressa na formação do sistema nervoso, mas já não é mais presente em outros tecidos.

Um outro viés de pesquisa seria acompanhar as mudanças nos padrões de expressão dos genes durante a indução do epiblasto. Embora no atual trabalho essas mudanças tenham sido analisadas antes e depois desse processo, seria relevante observar a expressão gênica em intervalos mais curtos, como a cada 8 horas. Dessa maneira, poderiam ser detectados padrões de aumento ou diminuição da expressão desses genes que estão diretamente correlacionados e que podem ocorrer em algum momento específico nessa janela de 48 horas de diferenciação.

Uma região até então não caracterizada e localizada mais próxima do gene *Sox2*, em relação à região SCR, também exibiu algumas características de região reguladora. Isso incluiu ligação a fatores de transcrição, modificações de histona e hipersensibilidade à nucleases. Embora a deleção dessa região não tenha resultado na ausência da proteína Sox2 durante os experimentos, seria interessante estudá-la mais a fundo para determinar se ela possui alguma influência sob a expressão de *Sox2*.

Por fim, a região SCR ainda é uma grande região genômica, com cerca de 8 mil pares de bases. Há evidências da presença de duas microrregiões dentro da SCR que exibem uma maior atividade reguladora. Dessa maneira, a região SCR poderia ser desmembrada em regiões menores de modo a melhor entender seu funcionamento e suas interações com diferentes fatores celulares. Isso poderia ser feito através da deleção de trechos diferentes dentro da região SCR para determinar quais porções dessa região possuem maior atividade reguladora.

REFERÊNCIAS BIBLIOGRÁFICAS

ACAMPORA, D.; GIOVANNANTONIO, L. G.; SIMEONE, A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. **Development**, Naples, v. 140, n. 1, p.43-55, 15 nov. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23154415>>. Acesso em: 04 out. 2015.

ALBERTS, B., *et al.* Biologia molecular da célula. 5. ed. Porto Alegre: Artes Médicas, 2010. ISBN 9788536320663.

BANERJI, J.; RUSCONI, S.; SCHAFFNER, W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. **Cell**, Zürich, v. 27, n. 2, p.299-308, dez. 1981. Disponível em: <<http://api.elsevier.com/content/article/PII:009286748190413X?httpAccept=text/xml>>. Acesso em: 28 set. 2015.

BARBER, R. D., *et al.* GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. **Physiological Genomics**, München, v. 21, n. 3, p.389-395, 1 mar. 2005. Disponível em: <<http://physiolgenomics.physiology.org/content/21/3/389.long>>. Acesso em: 29 ago. 2015.

BARRANGOU, R., *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. **Science**, Madison, v. 315, n. 5819, p.1709-1712, 23 mar. 2007. Disponível em: <<http://www.sciencemag.org/content/315/5819/1709>>. Acesso em: 03 out. 2015.

BERNSTEIN, B. E., *et al.* Methylation of histone H3 Lys 4 in coding regions of active genes. **Proceedings Of The National Academy Of Sciences**, Cambridge, v. 99, n. 13, p.8695-8700, 11 jun. 2002. Disponível em: <<http://www.pnas.org/content/99/13/8695.full#ref-1>>. Acesso em: 28 set. 2015.

BIEDA, M., *et al.* Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. **Genome Research**, Davis, v. 16, n. 5, p.595-605, 1 maio 2006. Disponível em:

<<http://genome.cshlp.org/content/16/5/595.full>>. Acesso em: 25 out. 2015.

BOCK, C.; WUTZ, A. DNA methylation: a matter of culture. **Nature Structural Molecular Biology**, Cambridge, v. 20, n. 3, p.249-251, 5 mar. 2013. Disponível em: <<http://www.nature.com/nsmb/journal/v20/n3/full/nsmb.2531.html#/author-information>>. Acesso em: 20 out. 2015.

BOYER, L. A., *et al.* Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. **Cell**, Cambridge, v. 122, n. 6, p.947-956, set. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3006442/>>. Acesso em: 04 out. 2015.

BRONS, I. Gabrielle M., *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. **Nature**, Cambridge, v. 448, n. 7150, p.191-195, 27 jun. 2007. Disponível em: <<http://www.nature.com/nature/journal/v448/n7150/full/nature05950.html>>. Acesso em: 04 out. 2015.

BROUNS, S. J. J., *et al.* Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. **Science**, Wageningen, v. 321, n. 5891, p.960-964, 15 ago. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18703739>>. Acesso em: 04 out. 2015.

BUECKER, C., *et al.* Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. **Cell Stem Cell**, Stanford, v. 14, n. 6, p.838-853, jun. 2014. Disponível em: <[http://www.cell.com/cell-stem-cell/abstract/S1934-5909\(14\)00140-4](http://www.cell.com/cell-stem-cell/abstract/S1934-5909(14)00140-4)>. Acesso em: 29 ago. 2015.

CALO, E; WYSOCKA, J. Modification of Enhancer Chromatin: What, How, and Why?. **Molecular Cell**, Stanford, v. 49, n. 5, p.825-837, mar. 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1097276513001020>>. Acesso em: 25 out. 2015.

CHAMBERS, I.; TOMLINSON, S. R. The transcriptional foundation of pluripotency. **Development**, Edinburgh, v. 136, n. 14, p.2311-2322, 19 jun. 2009. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/19542351>>. Acesso em: 27 set. 2015.

CHEN, X., *et al.* Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. **Cell**, Singapore, v. 133, n. 6, p.1106-1117, jun. 2008. Disponível em: <<http://api.elsevier.com/content/article/PII:S009286740800617X?httpAccept=text/xml>>. Acesso em: 13 set. 2015.

CREYGHTON, M. P., *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. **Proceedings Of The National Academy Of Sciences**, Cambridge, v. 107, n. 50, p.21931-21936, 24 nov. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3003124/>>. Acesso em: 28 set. 2015.

DAN, J., *et al.* Roles for Tbx3 in regulation of two-cell state and telomere elongation in mouse ES cells. **Scientific Reports**, Tampa, v. 3, n. 1, p.1-9, 13 dez. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3861804/#b5>>. Acesso em: 24 out. 2015.

DEB, K.; REESE, J.; PARIA, B. C. Methodologies to Study Implantation in Mice. In: SOARES, M. J.; HUNT, J. S. **Placenta and Trophoblast: Methods and Protocols**. Totowa: Humana Press Inc, 2006. p. 9-34. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16251731>>. Acesso em: 31 out. 2015.

D'HAESELEER, P. What are DNA sequence motifs? **Nature Biotechnology**, v. 24, n. 4, p.423-425, abr. 2006. Disponível em: <<http://www.nature.com/nbt/journal/v24/n4/full/nbt0406-423.html>>. Acesso em: 16 set. 2015.

EVANS, M. J.; KAUFMAN, M. H. Establishment in culture of pluripotential cells from mouse embryos. **Nature**, Cambridge, v. 292, n. 5819, p.154-156, 9 jul. 1981. Disponível em: <<http://www.nature.com/nature/journal/v292/n5819/abs/292154a0.html>>. Acesso em: 23 set. 2015.

FELSENFELD, G., *et al.* Chromatin structure and gene expression. **National Academy Of Sciences Of The United States Of America**, Bethesda, v. 18, n. 93, p.9384-9388, 3 set. 1996. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC38436/>>. Acesso em: 03 out. 2015.

FERRI, A. L. M., *et al.* Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. **Development**, [s.l.], v. 131, n. 15, p.3805-3819, 23 jun. 2004. Disponível em: <<http://dev.biologists.org/content/131/15/3805.long>>. Acesso em: 16 jun. 2015.

GARDNER, R. L. Contributions of blastocyst micromanipulation to the study of mammalian development. **Bioessays**, Oxford, v. 20, n. 2, p.168-180, 1998. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1521-1878\(199802\)20:23.0.CO;2-P/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1521-1878(199802)20:23.0.CO;2-P/abstract)>. Acesso em: 27 set. 2015.

GREBER, B., *et al.* Conserved and Divergent Roles of FGF Signaling in Mouse Epiblast Stem Cells and Human Embryonic Stem Cells. **Cell Stem Cell**, Münster, v. 6, n. 3, p.215-226, mar. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20207225>>. Acesso em: 27 set. 2015.

GRIFFITHS, *et al.* **Introdução à genética**. 9. ed. Rio de Janeiro: Guanabara Koogan, 2009. ISBN 9788527714976.

GROSS, D. S; GARRARD, W. T. Nuclease Hypersensitive Sites in Chromatin. **Annual Reviews Biochemistry**., Dalias, v. 57, n. 1, p.159-197, jun. 1988. Disponível em: <<http://www.annualreviews.org/doi/abs/10.1146/annurev.bi.57.070188.001111>>. Acesso em: 03 out. 2015.

HARDY, K. M., *et al.* FGF signalling through RAS/MAPK and PI3K pathways regulates cell movement and gene expression in the chicken primitive streak without affecting E-cadherin expression. **Bmc Developmental Biology**, Tucson, v. 11, n. 1, p.20-37, 2011. Disponível em: <<http://www.biomedcentral.com/1471-213X/11/20>>. Acesso em: 24 out. 2015.

HAYASHI, K., *et al.* Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. **Cell**, Kyoto, v. 146, n. 4, p.519-532, ago. 2011. Disponível em: <[http://www.cell.com/abstract/S0092-8674\(11\)00771-9](http://www.cell.com/abstract/S0092-8674(11)00771-9)>. Acesso em: 29 ago. 2015.

HEINTZMAN, N. D., *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. **Nature Genetics**, Boston, v. 39, n. 3, p.311-318, 4 fev. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17277777>>. Acesso em: 28 set. 2015.

HILLMAN, N.; SHERMAN, M. I.; GRAHAM, Chris. The effect of spatial arrangement on cell determination during mouse development. **Development**, Brighton, v. 28, n. 2, p.263-278, out. 1972. Disponível em: <<http://dev.biologists.org/content/28/2/263.abstract>>. Acesso em: 27 set. 2015.

IVANOVA, N., *et al.* Dissecting self-renewal in stem cells with RNA interference. **Nature**, Princeton, v. 442, n. 7102, p.533-538, 11 jun. 2006. Disponível em: <<http://www.nature.com/nature/journal/v442/n7102/abs/nature04915.html>>. Acesso em: 04 out. 2015.

IWAFUCHI-DOI, M., *et al.* Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. **Development**, Osaka, v. 139, n. 21, p.3926-3937, 19 set. 2012. Disponível em: <<http://dev.biologists.org/content/139/21/3926.article-info>>. Acesso em: 24 out. 2015.

JIANG, J., *et al.* A core Klf circuitry regulates self-renewal of embryonic stem cells. **Nature Cell Biology**, Singapore, v. 10, n. 3, p.353-360, 10 fev. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18264089>>. Acesso em: 04 out. 2015.

JOHNSON, D. S., *et al.* Genome-Wide Mapping of in Vivo Protein-DNA Interactions. **Science**, Stanford, v. 316, n. 5830, p.1497-1502, 8 jun. 2007. Disponível em:

<<http://www.sciencemag.org/content/316/5830/1497.figures-only>>. Acesso em: 03 out. 2015.

JOO, J. Y., *et al.* Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. **Scientific Reports**, Münster, v. 4, p.7477-7485, 17 dez. 2014. Disponível em: <<http://www.nature.com/articles/srep07477>>. Acesso em: 25 out. 2015.

KARMODIYA, K., *et al.* H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. **Bmc Genomics**, Strasbourg, v. 13, n. 1, p.424-442, 2012. Disponível em: <<http://www.biomedcentral.com/1471-2164/13/424>>. Acesso em: 28 set. 2015.

KENT, W. J., *et al.* The Human Genome Browser at UCSC. **Genome Research**, Santa Cruz, v. 12, n. 6, p.996-1006, 16 maio 2002. Disponível em: <<http://genome.cshlp.org/content/12/6/996.long>>. Acesso em: 29 ago. 2015.

KIM, J., *et al.* An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. **Cell**, Boston, v. 132, n. 6, p.1049-1061, mar. 2008. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0092867408003280>>. Acesso em: 25 out. 2015.

KOCH, F.; ANDRAU, J. Initiating RNA Polymerase II and TIPs as hallmarks of enhancer activity and tissue-specificity. **Transcription**, Marseille, v. 2, n. 6, p.263-268, nov. 2011. Disponível em: <<http://www.tandfonline.com/doi/abs/10.4161/trns.2.6.18747>>. Acesso em: 25 out. 2015.

KOPP, J. L., *et al.* Small Increases in the Level of Sox2 Trigger the Differentiation of Mouse Embryonic Stem Cells. **Stem Cells**, Omaha, v. 26, n. 4, p.903-911, abr. 2008. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1634/stemcells.2007-0951>>. Acesso em: 04 set. 2015.

KOUZARIDES, T. Chromatin Modifications and Their Function. **Cell**, Cambridge, v. 128, n. 4, p.693-705, fev. 2007. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0092867407001845>>. Acesso em: 28 set. 2015.

LATCHMAN, D. S. Transcription factors: An overview. **The International Journal Of Biochemistry & Cell Biology**, London, v. 29, n. 12, p.1305-1312, dez. 1997. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S135727259700085X>>. Acesso em: 03 out. 2015.

LEVINGS, P. P.; BUNGERT, J. The human β -globin locus control region. **European Journal Of Biochemistry**, Gainesville, v. 269, n. 6, p.1589-1599, 14 mar. 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11895428>>. Acesso em: 04 out. 2015.

LI, J.-y., *et al.* Synergistic Function of DNA Methyltransferases Dnmt3a and Dnmt3b in the Methylation of Oct4 and Nanog. **Molecular And Cellular Biology**, Shanghai, v. 27, n. 24, p.8748-8759, 15 out. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2169413/>>. Acesso em: 24 out. 2015.

LI, Q. Locus control regions. **Blood**, Seattle, v. 100, n. 9, p.3077-3086, 21 jun. 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2811695/>>. Acesso em: 25 out. 2015.

LI, Y., *et al.* CRISPR Reveals a Distal Super-Enhancer Required for Sox2 Expression in Mouse Embryonic Stem Cells. **Plos One**, San Diego, v. 9, n. 12, p.1-17, 8 dez. 2014. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114485>>. Acesso em: 23 out. 2015.

LINDSLEY, R. C., *et al.* Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. **Development**, St Louis, v. 133, n. 19, p.3787-3798, out. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16943279>>. Acesso em: 04 out. 2015.

LIU, P., *et al.* Requirement for Wnt3 in vertebrate axis formation. **Nature Genetics**, Houston, v. 22, n. 4, p.361-365, 1 ago. 1999.

Disponível em:

<http://www.nature.com/ng/journal/v22/n4/full/ng0899_361.html>.

Acesso em: 04 out. 2015.

MARTIN, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. **Developmental biology**, São Francisco, v. 78, n. 12, p.7634-7638, dez. 1981. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC349323/>>. Acesso em: 23 set. 2015.

MARUYAMA, T., *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. **Nature**, Massachusetts, v. 1, n. 33, p.538-542, 23 mar. 2015. Disponível em: <<http://www.nature.com/nbt/journal/v33/n5/full/nbt.3190.html>>. Acesso em: 25 out. 2015.

MASUI, S., *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. **Nature Cell Biology**, Kobe, v. 9, n. 6, p.625-635, 21 maio 2007. Disponível em: <<http://www.nature.com/ncb/journal/v9/n6/abs/ncb1589.html>>. Acesso em: 04 out. 2015.

MATHELIER, A., *et al.* JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. **Nucleic Acids Research**, Vancouver, v. 42, n. 1, p.142-147, 4 nov. 2013. Disponível em: <<http://nar.oxfordjournals.org/content/early/2013/11/04/nar.gkt997.full>>. Acesso em: 19 set. 2015.

MITCHELL, P., TJIAN, R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. **Science**, Berkeley, v. 245, n. 4916, p.371-378, 28 jul. 1989. Disponível em: <<http://www.sciencemag.org/content/245/4916/371.long>>. Acesso em: 03 out. 2015.

MOJICA, F. J. M., *et al.* Short motif sequences determine the targets of the prokaryotic CRISPR defence system. **Microbiology**, Alicante, v. 155, n. 3, p.733-740, 1 mar. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19246744>>. Acesso em: 04 out. 2015.

MOORE, K. L.; PERSAUD, T. V. N. **Embriologia clinica**. 8. ed. Rio de Janeiro: Elsevier, 2008. ISBN 9788535226621

Mouse ENCODE Consortium *et al.* An encyclopedia of mouse DNA elements (Mouse ENCODE). **Genome Biology**, Seattle, v. 13, n. 8, p.418-421, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3491367/>>. Acesso em: 13 set. 2015.

NATIONAL INSTITUTES OF HEALTH (NIH). (Org.). **What are the potential uses of human stem cells and the obstacles that must be overcome before these potential uses will be realized?** Disponível em: <<http://stemcells.nih.gov/info/basics/pages/basics6.aspx>>. Acesso em: 23 out. 2015.

NIWA, H., *et al.* Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. **Genes & Development**, Edinburgh, Scotland, v. 12, n. 13, p.2048-2060, 1 jul. 1998. Disponível em: <http://genesdev.cshlp.org/content/12/13/2048?ijkey=a53d3b91b1167f983772c6b13558d9f38449d9ff&keytype=tf_ipsecsha>. Acesso em: 25 out. 2015.

NIWA, H. How is pluripotency determined and maintained? **Development**, Kobe, v. 134, n. 4, p.635-646, 10 jan. 2007. Disponível em: <<http://dev.biologists.org/content/134/4/635>>. Acesso em: 23 set. 2015.

OKANO, M., *et al.* DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. **Cell**, Cambridge, v. 99, n. 3, p.247-257, out. 1999. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0092867400816566>>. Acesso em: 04 out. 2015.

PHILLIPS, J. E.; CORCES, V. G. CTCF: Master Weaver of the Genome. **Cell**, Atlanta, v. 137, n. 7, p.1194-1211, jun. 2009. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0092867409006990>>. Acesso em: 25 out. 2015.

POH, Y., *et al.* Generation of organized germ layers from a single mouse embryonic stem cell. **Nature Communications**, Urbana, v. 5, n. 4000, p.1-12, 30 maio 2014. Disponível em: <<http://www.nature.com/ncomms/2014/140530/ncomms5000/full/ncomms5000.html#ref6>>. Acesso em: 23 set. 2015.

ROSSANT, J. Stem Cells and Early Lineage Development. **Cell**, Toronto, v. 132, n. 4, p.527-531, fev. 2008. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0092867408001402>>. Acesso em: 23 set. 2015.

SANCHEZ-CASTILLO, M., *et al.* CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. **Nucleic Acids Research**, Cambridge, v. 43, n. 1, p.1117-1123, 30 set. 2014. Disponível em: <<http://nar.oxfordjournals.org/content/early/2014/09/30/nar.gku895.abstr>>. Acesso em: 19 set. 2015.

SANDER, J. D; JOUNG, J K. CRISPR-Cas systems for editing, regulating and targeting genomes. **Nature Biotechnology**, Massachusetts, v. 32, n. 4, p.347-355, 2 mar. 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24584096>>. Acesso em: 04 out. 2015.

SHLYUEVA, D.; STAMPFEL, G.; STARK, A. Transcriptional enhancers: from properties to genome-wide predictions. **Nature Reviews Genetics**, Vienna, v. 15, n. 4, p.272-286, 11 mar. 2014. Disponível em: <<http://www.nature.com/nrg/journal/v15/n4/full/nrg3682.html>>. Acesso em: 03 out. 2015.

SMALE, S. T.; KADONAGA, J. T. The RNA polymerase II core promoter. **Annual Reviews Biochemistry**, Los Angeles, v. 72, n. 1, p.449-479, jun. 2003. Disponível em: <<http://www.annualreviews.org/doi/abs/10.1146/annurev.biochem.72.121801.161520>>. Acesso em: 03 out. 2015.

SPITZ, F.; FURLONG, E. E. M. Transcription factors: from enhancer binding to developmental control. **Nature Reviews Genetics**, Heidelberg, v. 13, n. 9, p.613-626, 7 ago. 2012. Disponível em:

<<http://www.nature.com/nrg/journal/v13/n9/full/nrg3207.html>>. Acesso em: 25 out. 2015.

STRAHL, B. D.; ALLIS, C. D. The language of covalent histone modifications. **Nature**, Charlottesville, v. 403, n. 6765, p.41-45, 6 jan. 2000. Disponível em: <<http://www.nature.com/nature/journal/v403/n6765/full/403041a0.html>>. Acesso em: 28 set. 2015.

TAKAHASHI, K.; YAMANAKA, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. **Cell**, Kyoto, v. 126, n. 4, p.663-676, ago. 2006. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0092867406009767>>. Acesso em: 27 set. 2015.

TESAR, P. J., *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. **Nature**, Maryland, v. 448, n. 7150, p.196-199, 27 jun. 2007. Disponível em: <<http://www.nature.com/nature/journal/v448/n7150/abs/nature05972.html>>. Acesso em: 05 out. 2015.

THOMSON, J. A., *et al.* Embryonic Stem Cell Lines Derived from Human Blastocysts. **Science**, Madison, v. 282, n. 5391, p.1145-1147, 6 nov. 1998. Disponível em: <<http://www.sciencemag.org/content/282/5391/1145.full>>. Acesso em: 27 set. 2015.

THOMSON, M., *et al.* Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. **Cell**, Cambridge, v. 145, n. 6, p.875-889, jun. 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21663792>>. Acesso em: 04 out. 2015.

TOLHUIS, B., *et al.* Looping and Interaction between Hypersensitive Sites in the Active β -globin Locus. **Molecular Cell**, Rotterdam, v. 10, n. 6, p.1453-1465, dez. 2002. Disponível em: <[http://dx.doi.org/10.1016/S1097-2765\(02\)00781-5](http://dx.doi.org/10.1016/S1097-2765(02)00781-5)>. Acesso em: 04 set. 2015.

TOMIOKA, M., *et al.* Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. **Nucleic Acids Research**,

Saitama, v. 30, n. 14, p.3202-3213, 15 jul. 2002. Disponível em: <<http://nar.oxfordjournals.org/content/30/14/3202.long>>. Acesso em: 04 set. 2015.

UNTERGASSER, A., *et al.* Primer3Plus, an enhanced web interface to Primer3. **Nucleic Acids Research**, Wageningen, v. 35, p.71-74, 8 maio 2007. Disponível em: <http://nar.oxfordjournals.org/content/35/suppl_2/W71.long>. Acesso em: 29 ago. 2015.

WIEDENHEFT, B.; STERNBERG, S. H.; DOUDNA, J. A. RNA-guided genetic silencing systems in bacteria and archaea. **Nature**, Maryland, v. 482, n. 7385, p.331-338, 15 fev. 2012. Disponível em: <<http://www.nature.com/nature/journal/v482/n7385/full/nature10886.html>>. Acesso em: 04 out. 2015.

WILCOX, A. J.; BAIRD, D. D.; WEINBERG, C. R. Time of Implantation of the Conceptus and Loss of Pregnancy. **New England Journal Of Medicine**, North Carolina, v. 340, n. 23, p.1796-1799, 10 jun. 1999. Disponível em: <<http://www.nejm.org/doi/full/10.1056/NEJM199906103402304#t=articleTop>>. Acesso em: 31 out. 2015.

WOLPERT, L. **Princípios de biologia do desenvolvimento**. 3. ed. Porto Alegre Artmed, 2008. ISBN 978853631335.

YE, S., *et al.* Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. **Embo J**, Los Angeles, v. 32, n. 19, p.2548-2560, 13 ago. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3791365/>>. Acesso em: 25 out. 2015.

YING, Q.; SMITH, A. G. Defined Conditions for Neural Commitment and Differentiation. **Methods In Enzymology**, Edinburgh, Scotland, p.327-341, 2003.

YING, Q., *et al.* BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. **Cell**, Edinburgh, Scotland, v. 115, n. 3, p.281-292, out. 2003. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S009286740300847X>>. Acesso em: 25 out. 2015.

YING, Q., *et al.* The ground state of embryonic stem cell self-renewal. **Nature**, Los Angeles, v. 453, n. 7194, p.519-523, 22 maio 2008.

Disponível em:

<<http://www.nature.com/nature/journal/v453/n7194/abs/nature06968.html>>. Acesso em: 29 ago. 2015.

ZHANG, X., *et al.* Esrrb Activates Oct4 Transcription and Sustains Self-renewal and Pluripotency in Embryonic Stem Cells. **Journal Of Biological Chemistry**, Beijing, v. 283, n. 51, p.35825-35833, 28 out. 2008. Disponível em:

<<http://www.jbc.org/content/283/51/35825.full#ref-8>>. Acesso em: 04 out. 2015.

ZHANG, S.; KUHN, J. R. **Cell isolation and culture**. 2013. Disponível em:

<http://www.wormbook.org/chapters/www_cellculture/cellculture.html>. Acesso em: 13 set. 2015.

ZHOU, H. Y., *et al.* A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. **Genes Development**, Toronto, v. 28, n. 24, p.2699-2711, 15 dez. 2014. Disponível em:

<<http://genesdev.cshlp.org/content/28/24/2699.short>>. Acesso em: 29 ago. 2015.

ZHU, Q., *et al.* The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. **Elife**, Shanghai, v. 3, n. 1, p.1-21, 14 jun. 2014. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/24929964>>. Acesso em: 04 out. 2015.

ANEXO A

Sequências genômicas em torno das deleções realizadas pelo sistema CRISPR/Cas9. O símbolo “|” indica o ponto de quebra da dupla fita de DNA, enquanto a sequência sublinhada indica o alvo da enzima *Cas9* (ZHOU *et al.*, 2014).

Região Reguladora	Sequência genômica restante
ΔSRR1	CATATCATATAAGCCGTGGGACGAAATCATAAGAA TCAAATTTAATTTGCTTAAGTTAGAAATGTACATAG AACCGCCAGATAAGTGGGAGGTTAAGCGAGGGGA AAATAGGTATAAATTATGAGATTCAAAGTATGAAT GAATTATGATTGTCCGCGCGCCAC CTGCGGAATGC AGGGTGGTCTGCCCAGTGACCATTGTGCTCAATTCA TTCCACCTCTCACTGATGAAGT
ΔSRR2	ACAGTTTGAGATAAATAAATTTTTGAAATATGGAC ACTGAAATTATGCCTGAGTCTTTCATTTCTTTGTGG CTAGCATGCTTATACATAACTTCTTGGGGAAGTGA TAACTGTTTACGCCACCAAGATCTTGATGTTACCAGC AATGTCTGTTAGCTTTAACTGCAGAAAGAAAAAAT GGCCGAATGATTAATAACGGGGAAGTGAATGCGT TGACATTTCTAGAATAGAACTAGAAGCATATATT GTAAGTTCAC AGAGGCAGTGTTTCTAGTTGGCAGG CAAGTCAGGTGATTCAAACCTCGGAGGGGCCATTA GTTTTATGACCCCATCTGCTTGTCTATAGCGAGTCA AGTGGACACGGATGTATAAAGATTTTCGGCGCTTTG AACCCTTTTGCTTCTCCAGGCCAGGTTTCGG
ΔSRR18	GGGTCTGTGGCCTTCCTTTTACAGGAGAATAGATTG TCTTTCCTCTGTGGGAATGGGGCAAGTGAAAGAG AGGTGAAAAGACGGCTTTTGGGAGCTGGATAGCCC GGTCAATCTCCCAC TTTAAATGGCGC CTGTGGATCA TGATCATGGATGAACTCTGCTGTTT
ΔSCR (ΔSRR104-112)	CCCCTGGAAAAGCAGTAAGTGCTCTAACCACTGAG CTACCGTTC GAGTGTTCGAGAACAGTCAGGGATGC ACAGAGAAACCGTCTCAAAGACAAATACAATAAC CAAGACCAAA

Fonte: ZHOU *et al.* (2014)